

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:45:20 ; Search time 36.6935 Seconds
(Without alignments)
3043.156 Million cell updates/sec

Title: US-09-927-315-9

Perfect score: 4443
Sequence: 1 MGPRACTICLFLWLWLAFL.....ERNTPAYFNSMIGGYTMRD 838

Scoring table:
BIOSW62
Gapop 10.0 , Gapext 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: /SID52/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID52/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID52/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID52/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID52/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID52/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID52/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID52/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID52/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID52/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID52/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID52/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID52/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID52/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID52/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID52/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID52/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID52/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID52/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID52/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID52/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID52/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID52/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3514	79.1	669	21	AAV77558 Human GPCR-B4 poly
2	3203.5	72.1	843	21	AAV77556 Rat GPCR-B4 polype
3	3151.5	70.9	843	21	AAV77557 Mouse GPCR-B4 poly
4	1454	32.7	840	21	AAV45021 Rat sensory transd
5	1453	32.7	840	21	AAV45028 Rat sensory transd
6	1452	32.7	840	21	AAV45026 Rat sensory transd
7	1451	32.7	840	21	AAV45027 Rat sensory transd
8	1434	32.3	842	21	AAV45022 Rat sensory transd
9	1354	30.5	841	22	AAE10372 Mouse sensory tran
10	1354	30.5	841	22	AAE11969 Human taste recept
					Human novel G-prot

11	1354	30.5	841	23	ABW77319 Human G-protein co
12	1270	28.6	777	21	AAV45023 Human sensory tran
13	1267.5	28.5	1138	23	AAO21502 Human novel G-prot
14	1260.5	28.4	763	22	AAE11970 Chicken calcium-se
15	1107	24.9	1059	22	AAU00508 Dogfish shark kidn
16	1100	24.8	1026	18	AAW32059 Parathyroid calculi
17	1094	24.6	1078	19	AAW11889 Human parathyroid
18	1094	24.6	1078	19	AAW54846 Human parathyroid
19	1094	24.6	1078	19	AAW38274 Human parathyroid
20	1094	24.6	1078	20	AAV28840 Human calcium rece
21	1094	24.6	1078	20	AAV41780 Human parathyroid
22	1094	24.6	1078	20	AAW69565 Human parathyroid
23	1094	24.6	1078	21	AAV51827 Human calcium rece
24	1094	24.6	1078	21	AAV70325 Human wild type ca
25	1094	24.6	1078	22	AAW4391 Protein encoded by
26	1094	24.6	1078	22	AAW47822 HuCAR4.0. Homo sa
27	1093	24.6	1078	22	AAU02195 Cynomolgous monkey
28	1092.5	24.6	1027	23	ABW78761 Dogfish shark kidn
29	1092.5	24.6	1027	23	AAW60004 Shark kidney calculi
30	1088.5	24.5	1079	19	AAW54847 Rat kidney calcium
31	1088.5	24.5	1079	19	AAW38275 Rat kidney cell ca
32	1088.5	24.5	1079	20	AAV41781 Rat parathyroid ca
33	1088.5	24.5	1079	20	AAW64928 Rat kidney extrace
34	1088.5	24.5	1079	20	AAW69566 Rat parathyroid ca
35	1088.5	24.5	1079	21	AAV51828 Rat calcium recept
36	1088.5	24.5	1079	23	AAW47823 RakCAR3A. Rattus
37	1087.5	24.5	1085	19	AAW54844 Bovine parathyroid
38	1087.5	24.5	1085	19	AAW38272 Bovine parathyroid
39	1087.5	24.5	1085	20	AAV41778 Bovine parathyroid
40	1087.5	24.5	1085	20	AAW69563 Bovine parathyroid
41	1087.5	24.5	1085	21	AAV51825 Bovine calcium rec
42	1087.5	24.5	1085	23	AAW47820 BOPCAR1. Bovine s
43	1079	24.3	1088	17	AAW11888 Parathyroid calculi
44	1079	24.3	1088	19	AAW54845 Human parathyroid
45	1079	24.3	1088	20	AAV41779 Human parathyroid

ALIGNMENTS

RESULT 1
AAV77558 standard; Protein: 669 AA.
XX
AC AAV77558;
XX
DT 08-MAY-2000 (first entry)
XX
DE Human GPCR-B4 polypeptide.
XX
KW Sensory transduction G-protein coupled receptor; GPCR; GPCR-B4; human;
KW taste transduction pathway; taste receptor; foliate; fungiform; food;
KW Circumvallate; taste signaling; pharmaceutical.
XX
OS Homo sapiens.
XX
PN WO200006593-A1.
XX
PD 10-FEB-2000.
XX
PF 27-JUL-1999; 99WO-US17104.
XX
PR 28-JUL-1998; 98US-0095464.
XX
PR 17-DEC-1998; 98US-0112747.
XX
PI (REGC) UNIV CALIFORNIA.
XX
Zuker CS, Adler JE, Lindemeier J;
XX
WPI; 2000-195257/17.
XX
N-PSDB; AAZ58965.
XX
PT New isolated sensory transduction G-protein coupled receptor, useful


```

Db 121 DDLPLPLKDYSGYMPHHVAVIGPDNSESATVSNILSHFLIPQITYSALISDKLRKRF 180
QY 178 PALLRTTPSADHVEAMVOLMHPKRNMTIVYSSDTYGRDNGOLLGEVVAR-RDICIATF 236
Db 181 PSMLETRVPATHIEAMVOLMWHFQMNMTIVLVSDDDYGERNSHLSORLRTSDCIATF 240
QY 237 QETLPTLPONQMTSEEROLVTIYDKLOOSTARVVVSPDLTYHFENEYLRONFGTA 296
Db 241 QEVLPVPESQYMRSEQOQNDLIDKLKRTSARVVVSPDLTYHFENEYLRONFGTA 300
QY 297 VVIASESWAIDVNLNLTELGLGTFGLTITQSVPIPGSEFERMGPOAGPPLSRTSOS 356
Db 301 VVIASESWAIDVNLNLTELGLGTFGLTITQSVPIPGSEFERMGPOAGPPLSRTSOS 360
QY 357 YTCNODCDACNMNTSEFNILSLGGERVYVSAYAAVAHLHLRLHNGNRCTKQKY 420
Db 361 YTCNODCDACNMNTSEFNILSLGGERVYVSAYAAVAHLHLRLHNGNRCTKQKY 420
QY 417 PMQLLEIKVNFETLDHIFFDPOGVALHLEIVQMOMDRSONPFOVASYYPLOROLK 476
Db 421 PMQLLEIKVNFETLDHIFFDPOGVALHLEIVQMOMDRSONPFOVASYYPLOROLK 480
QY 477 NTODISMHTVNTIPMSCKRCQSGQKKPVGIVHCCFECIDCLPGTFNLHTEDEYEQ 536
Db 481 YINNVSWTPNNTVPIVMSCKRCQSGQKKPVGIVHCCFECIDCLPGTFNLHTEDEYEQ 540
QY 537 ACPNEMMSQSETEKROVLFEWHEARTIVALLAAGFISTALIVIFMRHQTPIV 596
Db 541 SCQGSWSTKNDITCFQRPRYLEWHEVPTIVALLAAGFISTALIVIFMRHQTPIV 600
QY 597 RSAGPMCFMLTLTLVAVVYVYVGPVKVSTCLCRQALFPLCEFTICISCIATVRSFOY 656
Db 601 RSAGPMCFMLTLTLVAVVYVYVGPVKVSTCLCRQALFPLCEFTICISCIATVRSFOY 660
QY 657 CAFKASRPRAVSTVWYQGPYVSMAFTVTKMVIIVIGMLARPOSH--RTDDDPKRI 714
Db 661 CEFKARRLPSAYSEFMRHGPYVAFITAKVALVAGNMILA-TTINIGRTDDDPRI 719
QY 715 TVSCNPNRNSLTFNTSLDILLVYGFSAVGMKELPNTYNAKRTILSMFTFTSSVS 774
Db 720 MTLSCPNRNSLTFNTSLDILLVYGFSAVGMKELPNTYNAKRTILSMFTFTSSVS 779
QY 775 LCTFNASVSGVLTIVDLVTVLNLATLSLGYFGPKCYMILFYPERNPAYNSMIQGYT 834
Db 780 LCTFNASVSGVLTIVDLVTVLNLATLSLGYFGPKCYMILFYPERNPAYNSMIQGYT 839
QY 835 MRR 837
Db 840 MRR 842

```

RESULT 3

AAV77557 standard; Protein; 843 AA.

AAV77557;

08-MAY-2000 (first entry)

Mouse GPCR-B4 polypeptide.

Sensory transduction G-protein coupled receptor: GPCR; GPCR-B4; mouse;
taste transduction pathway; taste receptor; foliate; fungiform; food;
circumvallate; taste signaling; pharmaceutical.

Mus sp.

WO200006593-A1.

10-FEB-2000.

27-JUL-1999; 99WO-US17104.

XX

```

PR 28-JUL-1998; 98US-0095464.
PR 17-DEC-1998; 98US-0112747.
PA (REGC ) UNIV CALIFORNIA.
PI Zuker CS, Adler JE, Lindemeyer J;
XX WPI: 2000-195257/17.
DR N-PSDB: AAZ38964..
XX
XX New isolated sensory transduction G-protein coupled receptor, useful
XX for developing products for use in studying and modulating the taste
XX transduction pathway and for generating taste topographic maps -
PS Claim 22; Page 69-70; 76pp; English.
CC
CC The invention provides nucleic acids encoding rat, mouse and human
CC sensory transduction G-protein coupled receptor (GPCR) polypeptides. The
CC GPCR polypeptides are components of the taste transduction pathway. The
CC nucleic acids can be used to identify taste cells and as tools for the
CC generation of taste topographic maps that elucidate the relationship
CC between the taste cells of the tongue and taste sensory neurons leading
CC to taste centers in the brain. GPCR-B4 is useful as a nucleic acid probe
CC for identifying subpopulations of taste receptor cells such as foliate,
CC fungiform, and circumvallate taste receptor cells. The polypeptides can
CC be used for identifying compounds that modulate sensory signaling in
CC sensory cells. Such modulators of taste transduction are useful for
CC pharmacological and genetic modulation of taste signaling pathways. These
CC modulatory compounds can then be used in the food and pharmaceutical
CC industries to customize taste. The present sequence represents a
CC mouse GPCR-B4 polypeptide.
SQ Sequence 843 AA;

```

Query Match 70.9%; Score 3151.5; DB 21; Length 843;
Best Local Similarity 68.9%; Pred. No. 7.6e-295;
Matches 581; Conservative 113; Mismatches 142; Indels 7; Gaps 4;

```

QY 1 MGPRAKTIQSLFFLLVYLAEP---AENSDFYLPQDYLGLSLNANMGYHNLQYLP 57
Db 1 MGPQARTLHLPLFLHLALPKPMLVGNSDPHLAGYLLGLFTLHANVKSVSHLSLOYLP 60
QY 58 MCKEYEVKIVGYNLQAMAFVEEINNDSSLPGVLYGEIVDVCYISNNQVPLYLELAH 117
Db 61 KQNEINMAYLGINLQAMAFVEEINNCSSLLPGVLYGEMDVCLSNINQDPLGLTFYSO 120
QY 118 EDNLPLIOEDYSNYISRVVAVIGPDNSESATVSNILSHFLIPQITYSALISDKLRKRF 177
Db 121 IDDFLPLIKDYSGYRQVAVVAVIGPDNSESATVSNILSHFLIPQITYSALISDKLRKRF 180
QY 178 PALLRTTPSADHVEAMVOLMHPKRNMTIVYSSDTYGRDNGOLLGEVVAR-DICIATF 236
Db 181 PSMLETRVPATHIEAMVOLMWHFQMNMTIVLVSDDDYGERNSHLSORLRTSDCIATF 240
QY 237 QETLPTLPONQMTSEEROLVTIYDKLOOSTARVVVSPDLTYHFENEYLRONFGTA 296
Db 241 QEVLPVPESQYMRSEQOQNDLIDKLKRTSARVVVSPDLTYHFENEYLRONFGTA 300
QY 297 VVIASESWAIDVNLNLTELGLGTFGLTITQSVPIPGSEFERMGPOAGPPLSRTSOS 356
Db 301 VVIASESWAIDVNLNLTELGLGTFGLTITQSVPIPGSEFERMGPOAGPPLSRTSOS 360
QY 357 YTCNODCDACNMNTSEFNILSLGGERVYVSAYAAVAHLHLRLHNGNRCTKQKY 420
Db 361 YTCNODCDACNMNTSEFNILSLGGERVYVSAYAAVAHLHLRLHNGNRCTKQKY 420
QY 417 PMQLLEIKVNFETLDHIFFDPOGVALHLEIVQMOMDRSONPFOVASYYPLOROLK 476
Db 421 PMQLLEIKVNFETLDHIFFDPOGVALHLEIVQMOMDRSONPFOVASYYPLOROLK 480
QY 477 NTODISMHTVNTIPMSCKRCQSGQKKPVGIVHCCFECIDCLPGTFNLHTEDEYEQ 536
Db 481 YINNVSWTPNNTVPIVMSCKRCQSGQKKPVGIVHCCFECIDCLPGTFNLHTEDEYEQ 540

```


CC sensory transduction. GPCR-B3 is specifically expressed in foliate
 CC and fungiform cells, with lower expression in circumvallate taste
 CC receptor cells of the tongue. The protein has an extracellular domain,
 CC seven transmembrane domains and an intracellular domain.
 CC The GPCR-B3 sequence is used to screen compounds that modulate sensory
 CC signalling in taste cells, especially taste modulators useful in
 CC pharmaceutical and food industries to customize taste. The sequence
 CC can also be used as probe for identifying taste cells and
 CC subsets of taste receptor cells such as foliate, fungiform and
 CC circumvallate. Such probes are also useful to generate taste
 CC topographic maps that elucidate the relationship between the taste
 CC cells of the tongue and sensory neurons leading to taste centres
 CC in the brain.
 CC Note: The present sequence is not given in the specification but is
 CC derived from rat GPCR-B3 sequence shown in page 75 (AA15021).
 CC
 XX Sequence 840 AA:
 Query Match 32.7%; Score 1451; DB 21; Length 840;
 Best Local Similarity 39.4%; Pred. No. 1,4e-130;
 Matches 324; Conservative 124; Mismatches 346; Indels 28; Gaps 11;
 QY 27 FYLGGDYLLGGLFSLHANKGIVHNLQVPMG-KEYEVKVIQYGLMAMFAYEINND 85
 DB 31 FSLPDEFLAGLFLSHGDCLOVRHRPL--VYSCDRSDSFNGHGYHLFQAMFVTEEDINNS 88
 QY 86 SSLPGLVGLGEIVDVCYSNNQPVLYEFLAHE-DNLPIQEDYSNTYSRVAVIYGPDS 144
 DB 89 SALLPNIITLGEIVDVCESANVYATLVLAQGRHIEIQRDLRNHSSKVAFFGPNT 148
 QY 145 ESMVTANFLSLFLPQITYSANIDELRDKVFPALITPSPADHVEAMVQLMIFRMN 204
 DB 149 DHAVTATALLPFLMPLVYSASSVYLSAKRRPFLRTVPSDRQVEMVQLQSPGMV 208
 QY 205 MIIIVYSSDVTGRNGQLGGERVARQDICIAPQELPFLQGNQNTSEKRLTIYVKL 264
 DB 209 WISLIGSTGDLGQLVQALELAVERGICVAFKIDVPE-----SARVGDPRQMSMOHL 262
 QY 265 QOSTARVVVVESSPDLITLHFENEVLRFONTGAVVIASSMAIDPVLHMLTELGHGTLG 324
 DB 263 AQARTTVVVFSSNRHLAVFRSVYLANLTKVWASDMLSTIYTVTGIGIGIYVLG 322
 QY 325 ITIOSVPIPGSEFERWGPQAGPPPLSRTSQTTC--NQECDNCLNATLSFNTILRLSGE 382
 DB 323 VAVOQRRVGLKEFEESYVRAVTAAPSACPSGMCSTNQLRCHETFTTRMPTLGARSM 382
 QY 383 RVVYSVYAVYAVNAHALSLGCDKSTCKRVYVPMOLEIMKYNFLLHOFDPDQG 442
 DB 383 SAATRVYEAIVAVAHGLHQLGCTSEISCKRGPVPMQLQIYKYNFLHENTVAFDNG 442
 QY 443 DVALHLEIVQWQMRQNPQSV--ASYYPLOROLKNIDISWHTVNTIMSMSCSKRQ 500
 DB 443 DTLGYYDIIADWNGPWEFTFETISASLSPVHDI-NKTKIOMHGKNNQVVSCTTCL 501
 QY 501 SGQKKRPVIGIVCCPECIDCLPGTFPLNTEDEYEQACPNMWSYQSTSCPKROLVLE 560
 DB 502 AGHRVYVGSCHCECEPCVCAEATFLNMSE-LHICQPCGTENMAPKESTTCFPRVERLA 560
 QY 561 WHEAPTIIVALLALGFLSTALILY-----IFMRHPTPIVRSAGGPMCLMTLLIYA 614
 DB 561 WHE--PISLVILAA---NTLLLLVGTAGLFAWHHTPVRVRSAGGLCTLMGLSLVAG 614
 QY 615 YMVVYVYVPPKPVSTCLQALFPLCFTICISIAVRSFQIVCAFKMARPRAYSVWR 674
 DB 615 SCSPFSFGEPYVACILRQPLFLSLGFAIFLSCLTIRSFOLVITFKSTYKPTFRMAQ 674
 QY 675 YQGPVSAFTVYLMKVIVYGMALARPOSHPRTDDBPKITIVSCNPNYRNSLLFNMSLD 734
 DB 675 NHGGLFIYVYSTHLLICTLWLMVMTPTPTREORFPHILILECTEVNSGFLIAFHN 734
 QY 735 ILLSYVGSFAYMGRELPNTNNEAKFTLSMTYFTSSVSLCTPMSAVSGLVNTIYDLY 794
 DB 735 ILLSTVCSYLGKELEPENTNEACVTFSLINLVSWIAFTMASTIYQSGSTLPAVNVLA 794

QY 795 TVNLNLAISLGYFGKCYMLFPERNPAYENSMIOGYTMR 836
 DB 795 GITTLSGGFSGYFLPKCYVILCRPELNTEHFQASTQDYYTR 836
 RESULT 8
 ID AA15022 standard; Protein: 842 AA.
 AA15022
 AA15022;
 31-MAY-2000 (first entry)
 DE Mouse sensory transduction G-protein coupled receptor-B3.
 XX
 XX Mouse; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;
 KW sensory cell; taste receptor cell; screen; taste modulator;
 KW pharmaceutical; food industry; taste topographic map; tongue.
 XX Mus sp.
 OS
 XX W0200006592-A1.
 XX
 XX 10-FEB-2000.
 XX
 XX 27-JUL-1999; 99WO-US17099.
 XX
 XX 28-JUL-1998; 98US-0094465.
 XX
 XX (REGC) UNIV CALIFORNIA.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 XX Zuker CS, Adler JE, Lindemeier J, Ryba N, Hoon M;
 DR WPI: 2000-205451/18.
 XX N-PSDB: AA150744.
 XX
 XX New isolated sensory transduction G-protein coupled receptor, useful
 PT for developing products for use in studying and modulating the taste
 PT transduction pathway -
 XX
 XX Claim 22; Page 75; 83pp; English.
 XX
 CC The present sequence is a taste cell specific G-protein
 CC coupled receptor, GPCR-B3 which is involved in sensory transduction.
 CC This sequence was obtained from mouse circumvallate and foliate papillae.
 CC GPCR-B3 is specifically expressed in foliate and fungiform cells, with
 CC lower expression in circumvallate taste receptor cells of the tongue.
 CC The present sequence is used to screen compounds that modulate sensory
 CC signalling in taste cells, especially taste modulators useful in
 CC pharmaceutical and food industries to customize taste. The sequence
 CC can also be used as probe for identifying taste cells and
 CC subsets of taste receptor cells such as foliate, fungiform and
 CC circumvallate. Such probes are also useful to generate taste
 CC topographic maps that elucidate the relationship between the taste
 CC cells of the tongue and sensory neurons leading to taste centres
 CC in the brain.
 CC
 XX Sequence 842 AA;
 SQ
 Query Match 32.3%; Score 1434; DB 21; Length 842;
 Best Local Similarity 39.6%; Pred. No. 6.2e-129;
 Matches 327; Conservative 123; Mismatches 340; Indels 36; Gaps 15;
 QY 27 FYLGGDYLLGGLFSLHANKGIVHNLQVPMG-KEYEVKVIQYGLMAMFAYEINND 85
 DB 33 FSLPDEFLAGLFLSHGDCLOVRHRPL--VYSCDRSDSFNGHGYHLFQAMFVTEEDINNS 90
 QY 86 SSLPGLVGLGEIVDVCYSNNQPVLYEFLAHE-DNLPIQEDYSNTYSRVAVIYGPDS 144
 DB 91 TALPNIITLGEIVDVCESANVYATLVLAQGRHIEIQRDLRNHSSKVAFFGPNT 150

[illegible]

RESULT 9	
AAE10372	
ID	AAE10372 standard; Protein; 841 AA.
XX	
XX	
AC	AAE10372;
XX	
DT	10-DEC-2001 (first entry)
XX	
DE	Human taste receptor, hT1R1 protein.
XX	
KW	Human; taste-cell-specific G protein-coupled receptor; hT1R1; drug;
KW	genetic modulation; pharmaceutical; taste sensation; food industry;
XX	chemosensory transduction.
XX	
OS	Homo sapiens.
XX	
XX	
Key	Location/Qualifiers
FH	Misc-difference 274
FT	/note="Encoded by TTTT"
XX	
PN	WO200166563-A2.
XX	
PD	13-SEP-2001.
XX	

Query Match	Best Local Similarity	Matches	314; Conservative	131; Mismatches	355; Indels	40; Gaps	16;
07-MAR-2001; 2001IMO-US07265;							
07-MAR-2000; 2000US-0187546;							
07-APR-2000; 2000US-0195536;							
06-JUN-2000; 2000US-0209840;							
23-JUN-2000; 2000US-0214213;							
17-AUG-2000; 2000US-0226448;							
03-JAN-2001; 2001US-0259227;							
(SENO-) SENOMIX INC.							
Adler JE, Zozulya S, O'Connell SM, Li X, Staszewski L;							
WPI; 2001-582267/65.							
N-PSDB; AAD17515, AAD17516.							
New mammalian taste-cell-specific G protein-coupled receptor polypeptides for identifying compounds that modulate taste signaling are useful in food, to modulate the sweet taste of foods or drugs -							
Claim 152; Page 83; 119pp; English.							
The invention relates to mammalian taste-cell-specific G protein-coupled receptors, T1R and their corresponding cDNA molecules. Taste receptors, T1R are useful for screening compounds which are used to activate or modulate chemosensory transduction, such as taste sensation. The identification and isolation of novel taste receptors and taste signalling molecules allow for new methods of chemical and genetic modulation of taste transduction pathways. The taste modulating compounds are useful in pharmaceuticals and food industries to improve the taste of a variety of consumer products, or to block undesirable tastes, e.g., in certain pharmaceuticals. T1R's are also useful in biochemical assay for identifying tastant (T1R) ligands having binding specificity for T1R involved in taste signalling. The present sequence is human taste-cell-specific G protein-coupled receptor, hT1R1 protein.							
Sequence 841 AA;							
30.5%; Score 1354; DB 22; Length 841;							
Best Local Similarity 37.4%; Pred. No. 3.3e-121;							
Matches 314; Conservative 131; Mismatches 355; Indels 40; Gaps 16;							
16 WYLA -EPAENS -DFLPBGYLLGGLFSLHANKGYHLNLFLOYPMC -KEYEKVIGYNLM 72							
19 WAFACHSTSSDPFLPBGYLLAGLPLHSGCLQVKNRP -EYLLCDRCSENEHGYHLF 76							
73 QAMRAVEEINNDSILPGVLLGYEIVDCYISNNVPALYFLA---REDNLLPIQEDY 128							
77 QAMRGVEEINNSTALLPRITLGYQLDYDCDSANVATLRYLSLPGQHH---IELQGLD 133							
129 SNTISRVAVAVIGPDNSESVYVAFNLSLFLPQITYSAISDELROKVRPALIRTPSPAD 188							
134 LHYSPYLAVALVIGPDSTNRATTAALISPLVPMISVAASESLSVKROYSPSEFLRTPDNK 193							
189 HHVEAVOLMLHFRNNWMLIVLSPPYGDNQOLLGGERVARDICAFETPL-TLQPNQ 247							
194 YQVEHMLLLQKFGTWTWISLVGSSDDYGLQVQALENQATGGICIAFDIMPFSAQVED 253							
248 NMTSEORLVTYIVDKLOOSTARVVVYVSPDLLYHFEFNEVLRQNETGAWIASESAID 307							
254 -----ERMQLMRLHQAQATVYVVFSSRQLARVFEFSVVLTNLGRVWVASEAMALS 306							
308 PYLHMLTELGHGTLGITIQSVPLPGSEFEEMGPQ---GPPLSKTSQSTYCNQED 364							
307 RHITGVPGIORIGMVLGAIVOKRAVPGKAFEEAARADKKAPRCHKSMSC-SSNQLCR 365							
365 NCLNMTLSFNTLISGERVVVYVSAVYAVVAHALSHLSLGCRCSPCTKRVRVPMOLLEI 424							
366 ECOLAMATHMPLKAFKSMSSATNATRAYAYAAHGLHOLLGCGASGACSRKRYVPMOLLEI 425							
425 WKVNFETLLDHOIFEDPGQVVALHLEIVQWQMRSONPQSVAS--YPLQROLKNIQDIS 482							
426 HHVEHLLHDTYAFDNDNRPLSSYNIADWNGPKPTEFVVLGSSSPSPOLNI-NETKIQ 484							

OY		483	HHNNITIPMSKSKRQSQKKRPVGHVCOFECIDLPFTFLNHTDEVEQCAPNNE	542
Db		485	WHGKDNOVPKSVSSDCLBQHVVTFHFHCCPCVPGAGTFLNK-S-DLRCOPCGKEE	543
OY		543	WSYSEIISCKRKRLVFLEMEAPTIALVALAALGEFLSTAILIVIFWRHFOPIYRSAGP	602
Db		544	WAPSGTCPCPRVTVFLALEHN SWVLLANNTLLLLLLGTAGLEFWMLDTFPVRASGR	603
OY		603	MCFMLTLULLVAIVWVYVYGPPKYSTCICRALFPLCFTICISCIAVRSQIYCAFKA	662
Db		604	LCFMLTSLAAGSGSILGFEGEPTRPACLLRQALPALFTTFLSCITFRSQTIIIFKS	663
OY		663	SRFPAASYWRVGQPYVSMAFTTYVKVIYVIGMIARQSHPRDDPDPTITVSCNP	722
Db		664	TKVTFTHAVWQNNGAGLFEMISSAQOLICTLTVWTPLPARHYQRFFHLVLECTET	723
OY		723	YRNLSL-----LENTSIDLLSVGFSFAVMGKELPTYNNAKFTTSMTEFTTSYSLC	776
Db		724	--NSIGETLAFLFYNG----LTLSIAFACSYLKDPENVNNAKCVTESLLENFVSWIAFE	777
OY		777	TFMSAYSGVLYTIYIDLTVYLNLALSIGYREPKCYMILLFPERTPAYEFMSMQGYTM	836
Db		778	TTASVTDGYKTPAANMMAGSLSSGFGGYFLPKCYVILCRPDLNSTEHQAISTDYTR	837
RESULT		10		
ID	AAE11969	standard; protein; 841 AA.		
XX	AAE11969;			
XX				
DT	18-DEC-2001	(first entry)		
XX				
DE		Human novel G-protein coupled receptor (NGPCR) protein #1.		
KW		Human; G-protein coupled receptor; GPCR; gene therapy; drug screening;		
RW		antisense-therapy; signal transduction; behavioural disorder; obesity;		
KW		heartbeat rate; inflammation; immune disorder; diabetes; cancer;		
KW		coronary disease.		
XX				
OS	Homo sapiens.			
XX				
FH	Key	Location/Qualifiers		
FT	Misc-difference 107	/note= "Encoded by TYT"		
FT	Misc-difference 372	/note= "Encoded by RGA"		
XX				
PN	WO200172842-A2.			
XX				
PD	04-OCT-2001.			
XX				
PE	28-MAR-2001; 2001MO-US09996.			
XX				
ER	28-MAR-2000; 2000US-192978P.			
XX				
PA	(LEXI-) LEXICON GENETICS INC.			
PI	Malke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abulin A;			
PI	Zambrowski B, Sands AT,			
XX				
DR	WPI; 2001-616474/71.			
XX				
N-P	N-PDB; AAD19501.			
XX				
PT	Novel isolated polynucleotides encoding human G protein coupled			
PT	receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately			
XX	expressed NGPCRs for diagnosis of disease, and as probes or primers			
PS	Claim 3; Page 72-74; 80pp; English.			
CC	The present sequence is human novel G-protein coupled receptor			
CC	(NGPCR) protein. NGPCRs are transmembrane proteins that span			

DB 724 - -NSIGLIFLAFYNG- - -LLSISAFACSYLGDLPENYNKACVPEFLLNFVSWIAFE 777
 OY 777 TFSASVAGVLTIVDLVTLVNLALSLGEGPCWMLTYPERNRPAYFNSMIOGTMR 836
 DB 778 TFSASVAGVLTIVDLVTLVNLALSLGEGPCWMLTYPERNRPAYFNSMIOGTMR 837

RESULT 11
 ABB77319
 ID ABB77319 standard; Protein; 841 AA.
 AC ABB77319;
 XX
 DT 17-JUN-2002 (first entry)
 XX
 DE Human G-protein coupled receptor SRO ID NO 3.
 XX
 KW Human; GPCR; G-protein coupled receptor; receptor; anti-HIV; antitumour;
 KW antinflammatory; antiallergic; antianaemic; antiasthmatic; virucide;
 KW immunosuppressive; dermatological; nephrotoxic; antidiabetic;
 KW cytostatic; neuroprotective; osteopathic; antiproliferative; antineoplastic;
 KW antirheumatic; thyromimetic; antileukemia; ophthalmological; antidiabetic;
 KW fungicide; antiparasitic; protozoacide; antihelminthic; antidiabetic;
 KW antidiabetic; antidiabetic; hepatotropic; anticonvulsant; anorectic; metabolic;
 KW antiparkinsonian; depilatory; tranquilizer; hypotensive; vasotropic;
 KW cardiologic; antineoplastic; antineoplastic; antineoplastic; antineoplastic;
 KW Parkinson's disease; Alzheimer's disease; Huntington's disease;
 KW cardiovascular disorder; acquired immunodeficiency syndrome; AIDS;
 KW Crohn's disease; diabetes mellitus; rheumatoid arthritis.
 XX
 OS Homo sapiens.
 XX
 PN MO200196323-A2.
 XX
 PD 27-DEC-2001.
 XX
 PF 15-JUN-2001; 2001MO-US19354.
 XX
 PR 16-JUN-2000; 2000US-212483P.
 PR 23-JUN-2000; 2000US-213950P.
 PR 26-JUN-2000; 2000US-214062P.
 PR 07-JUL-2000; 2000US-216595P.
 PR 14-JUL-2000; 2000US-218936P.
 PR 19-JUL-2000; 2000US-219154P.
 XX
 PA (INCY-) INCYTE GENOMICS INC.
 XX
 PI Lal P, Graul R, Hafalia AJA, Walla NK, Thornton M, Nguyen DB,
 PI Lu Y, Gandhi AR, Patterson C, Kallilick DA, Baughn MR, Ramkumar J;
 PI Tribouley CM, Lee EA, Ding L, Burford N, Yao MG, Yang J;
 PI Griffin JA;
 XX
 DR WPI, 2002-139780/18.
 DR N-PDB; ABL55953.
 XX
 PT Novel G-protein coupled receptor protein and polynucleotides useful for
 PT diagnosing, treating or preventing disorders of cell proliferation e.g.
 PT cancer, neurological and genetic disorder e.g. thalassemia
 XX
 PS Claim 49; Page 107-109; 121pp; English.

CC condition associated with decreased expression or overexpression of
 CC functional GPCR. Examples of disorders include cell proliferative
 CC disorder such as arteriosclerosis, atherosclerosis, hepatitis, mixed
 CC connective tissue disease (MCTD), psoriasis and cancer including
 CC adenocarcinoma, leukemia; a neurological disorder such as epilepsy,
 CC stroke, Alzheimer's disease, Huntington's disease, dementia, Parkinson's
 CC disease, retinitis pigmentosa, multiple sclerosis, bacterial and viral
 CC meningitis, abscess, subdural empyema; prion disease including kuru,
 CC Creutzfeldt-Jakob disease; fatal familial insomnia; neurofibromatosis,
 CC tuberous sclerosis, cerebral palsy, polyomyelitis; inherited, metabolic,
 CC endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis;
 CC mental disorders including mood, anxiety, and schizophrenic disorders;
 CC seasonal affective disorder (SAD); akathisia, amnesia, cataplexy,
 CC diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses,
 CC Tourette's disorder; cardiovascular disorders such as hypertension,
 CC vasculitis, vascular tumours, congestive heart failure, ischemic heart
 CC disease, myocardial infarction, calcific aortic valve stenosis,
 CC infective endocarditis, endocarditis of systemic lupus erythematosus,
 CC cardiac transplantation; gastrointestinal disorder such as dysphagia,
 CC gastritis, anorexia, nausea, emesis, abdominal angina, infections of
 CC the intestinal tract, peptic ulcer, hepatitis, cirrhosis, diarrhoea;
 CC acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, Reye's
 CC syndrome, liver infarction; an autoimmune/inflammatory disorder such as
 CC AIDS, Addison's disease, adult respiratory distress syndrome,
 CC allergies, amyloidosis, anaemia, asthma, atherosclerosis, autoimmune
 CC thyroiditis, bronchitis, Crohn's disease, diabetes mellitus,
 CC Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis,
 CC multiple sclerosis, myasthenia gravis, myocardial or pericardial
 CC inflammation, osteoporosis, psoriasis, rheumatoid arthritis,
 CC scleroderma, Sjogren's syndrome, systemic lupus erythematosus, uveitis,
 CC viral, bacterial, fungal, parasitic, protozoal, and helminthic
 CC infections, and trauma; a metabolic disorders such as diabetes,
 CC obesity, and osteoporosis. The GPCR encoding polynucleotide is useful
 CC for gene therapy and for creating knock in humanised animals (pigs) or
 CC transgenic animals (mice or rats) to model human disease.
 XX
 SQ Sequence 841 AA;
 XX

Query Match 30.5%; Score 1354; DB 23; Length 841;
 Best local Similarity 37.4%; Pred. No. 3.3e-121;
 Matches 314; Conservative 131; Mismatches 355; Indels 40; Gaps 16;

OY 16 WTAA-EPAENS-DPYLPDYLGLGSLAAMKGIYHNLQVPMC-KEYEVYVIGTMT 72
 DB 19 WAFACSTESSPDFTLPGLDGLPLHSGCLQVHRP--EYTLDRSCFENEGHLE 76
 OY 73 QAMRANEEIINNDLSLPGVLGLEYIVDVCYINNQPVLXFLA---HEDNLPLOEDX 128
 DB 77 QAMRLGVEETINNSTALPNTLGLYDVCSDANVATATRLVSLPGQHH--TELQGL 133
 OY 129 SNYISRVAVIGPDSSESVTVANFLSLFLPOLITSAISDELDRKVRFPALRTTPSAD 188
 DB 134 LHSPTLVAVIGPDSSTRAATTAALSPFLVPMISVAASETISVRRQYSPFLRTIPDX 193
 OY 189 HAYEVQMLHFRMMNITVLSDDYGGDNGULGERVARRDICTAFOETLP-TIQPNO 247
 DB 194 YQVEITWVLLOKFGWTWISLVGSSDDYGGQVQVLALEQAGQICIAFKRIMPFSAQVGD 253
 OY 248 NMTSEERORLVTIVDKLOOSTARVYVSPDLTLHFHFNFLVNLQNTGAVWIMSEMAID 307
 DB 254 -----ERMOCLMRHLAOGATVYVVFSSROLARPFESVTLNLTGKVAVASEMAALS 306
 OY 308 PVLNLTGLHGLTFLGITITVSPVIRGFSFEREMGDA--GPPPLSRISQSYTCNOECD 364
 DB 307 RHITGVGIGRIGVAGLQKRAVGLKAFEEAVARADKRAPCKGSMC--SSNOLCR 365
 OY 365 NCLNATLSFTTIRLISERVYVYSAVYAVAAHLSLDCGKSTCKRVYVWOLLEI 424
 DB 366 ECAAFNAHTPKRKAFSSMSAVYAVYAVAGLTLQGLCAGACSGRGVYVWOLLEI 425
 OY 425 WKVNFETLLDQIFEDPGDVALHLEIYQWQMDSONPQSVAS--YYPLROLKNIDIS 482
 DB 426 HKVHFLHKTVAFNDRNDPLASTYNIIMDMNGPKMTFTVLSSTVSPVLNI-NETKIO 484

RESULT 13
 AAO21502
 ID AAO21502 standard; Protein: 1138 AA.
 XX
 AC AAO21502;
 XX
 DT 15-AUG-2002 (first entry)
 XX
 DE 1138-mer rat protein relating to the human GPCR.
 XX
 KM G-protein coupled receptor; GPCR; gene chip; rat; immune response;
 KM chromosomal position; transgenic animal; gene therapy.
 XX
 OS Rattus norvegicus.
 XX
 PN W0200230981-A1.
 XX
 PD 18-APR-2002.
 XX
 PF 13-MAR-2001; 2001WO-0507832.
 XX
 PR 10-OCT-2000; 2000US-0684393.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Wei M, Zhong W, Ketchum KA, DiFrancesco V, Beasley EM;
 XX WPI; 2002-444173/47.
 DR
 XX
 PT Novel G protein coupled receptor, useful for raising antibodies, to
 PT elicit immune response, and as a reagent in assays designed to
 PT quantitatively determine levels of protein in biological samples -
 PS Disclosure; Page 75-78; 82pp; English.
 XX
 CC The invention relates to an isolated G-protein coupled receptor (GPCR)
 CC polypeptide, comprising an 852 residue amino acid sequence, given in the
 CC specification, an allelic variant or ortholog of the protein, or a
 CC fragment comprising at least 10 contiguous amino acids of the protein.
 CC GPCR is useful for identifying a modulator of GPCR and an agent that
 CC binds to GPCR. GPCR and a gene chip comprising GPCR are useful as models
 CC for the development of human therapeutic agents. GPCR is useful for
 CC raising antibodies, to elicit immune response, as a reagent in assays
 CC designed to quantitatively determine levels of protein in biological
 CC samples, and as markers for tissues in which the corresponding protein
 CC is preferentially expressed. A gene chip containing GPCR is also useful
 CC as a probe for determining the chromosomal positions of nucleic acid
 CC molecules by means of in situ hybridisation, in making vectors containing
 CC the gene regulatory regions of a gene chip containing GPCR, for designing
 CC ribozymes, in making vectors that express GPCR, and for constructing host
 CC cells and transgenic animals expressing nucleic acid molecules and
 CC peptides. A host cell containing GPCR is useful for conducting cell-based
 CC assays involving GPCR protein or its fragments, and for identifying
 CC protein mutants in which these function is affected. The polynucleotide
 CC encoding GPCR can be used to treat disorders by gene therapy. This
 CC sequence represents a 1138-mer rat protein relating to the human G-
 CC protein coupled receptor of the invention.
 XX
 XX
 S0 sequence 1138 AA;
 Query Match 28.5%; Score 1267.5; DB 23; Length 1138;
 Best Local Similarity 28.8%; Pred. No. 1.2e-112;
 Matches 330; Conservative 127; Mismatches 331; Indels 359; Gaps 21;

Db 120 LPLGYDCESELLAGRIYALIGPSSALLPNITLGEYELDYCESANAYATRLVALAQPRH 179
 QY 122 LPTOEDYSNYSIRVANYIGDN--SESVMYANFLSLFLPOITYSA----- 166
 Db 180 IEIOKDLRNHSSKVVAFIGPDNOSELAMVTGKFFSFLMPOVSTGASMELLSARETPPSF 239
 QY 167 ----- 166
 Db 240 FRIVPSDRVOLTAABELDFEGNATFLMPYASLSAFSPFRVPSDRQLLRGWSIDHA 299
 QY 167 -----ISDE-----LRDKVRPALIRTPPSADHVEANVQLMHPRW-NMT 206
 Db 300 VTTAALGPLMLPLVSYEASVYLSAKRKPSFLRTVPSDRHOVEVAVQLISFGQNMNV 359
 QY 207 IYVSSDPTGNDNGQLGGERVARDCIARQETLPTLOPNQNMTE----- 252
 Db 360 AALGSDEYGRQGLSTFSALAAARGICTAHGIVPLPRADSRILKQVDYLVHOVNMGSY 419
 QY 253 -----ERORLVTTI-----VDKL----- 264
 Db 420 GGLARGICAVPRDRQGSVWISLIGSYDQGLQVQALAEELAVPRGICVAFKDIYPSARV 479
 QY 265 -----QOSTARVYVF-----SPD----- 278
 Db 480 GDPROMQMOHIAQOSSVOVLLFASVAAHALFNYSISSRLSPKVVASEAWLTSOLVM 539
 QY 279 -----LPLX----- 289
 Db 540 GLPMAOMGVYLGFLVYFHAFLKVVASSEMSGGVLCARTVVVVFNRHLARFFESV 599
 QY 290 RQNFYGAIVIASSEWALDPLVNLTELHGTFLGITQSVPLPGSEFRMGQACPP 349
 Db 600 LANLTGKVVWASEDMALSTYTSVIGIGTVGVAVQROVGLKEFESESVRAVTA 659
 QY 350 LSRISQSYTC--NQCNCNLATISFNTILRLSERVYYSVAVARHSLGCDK 407
 Db 660 PSACPESQSTNOLCRECHTFTTRNMTLGAFSMAARYEAVYAVAGLHQLGCTIS 719
 QY 408 STCRKRVYPMOLLEETKRVFTLLDQIFPDPOGVALHLEIYQOMQDRSQNFQSV-- 465
 Db 720 EICRGVYPMQLQOYIKVNFLEHNTVAFDDNGDLGYDIAMMNGBEWTFEILIGS 779
 QY 466 ASYVPLORQKINIDISMTVNTIPMSKSCRCQSGQKKRPVGIHYCCFECIDLPGETF 525
 Db 780 ASLSPVHLDI-NKTKIQMHGKNQVPVSVCTDCLAGHRRVVGSHHCCEVCCEAGTE 838
 QY 526 LNHTEDEYEQACNNEMWSQSESCFRQOLVLEHNEAPRIANALLAALGELSTALIV 585
 Db 839 LNMSE-LHIQOPCTEEMAPBESTTCCPRVVEFLAMHE--PISLVLTAA---NTLLLL 891
 QY 586 -----IFMRFORPIYRSAGGPMCFMLTLTLVAVWVPVYVPGVSTCLQALPPL 639
 Db 892 LVYAGLIFAMHFTPVRSAGRLCFMLGSLVAGSCSFSEFEPVPACLRQPLPSL 951
 QY 640 CFTICISIAVRSFQIACFMASSRPFRAYSVYVRYQGPVYNAFITYLVKVIYVIGMA 699
 Db 952 GFALFELSCILTRSFQVLIIFEFSTKVPLEFRYMAQNNAGLFLVSSVHLLCFLWLM 1011
 QY 700 RPOSHPRTPDDPKITTVSCNPNTRNSLLENTSLDLISVYGSEFAYMGELPNTYEA 759
 Db 1012 WTPPRTEYORFPHVILIECTEVNSVGLFLAFTHNILLISTFVCSYLGKELPNTYEA 1071
 QY 760 FTLTSMFTFSSVSLCTCPMSAVSGLVYTYDILVYVNLALISLGEFGRCVYILTYPE 819
 Db 1072 CVYFSLILNYSWIAFTPMASTYIGSYLPAVNAVLAGLTLISGSGFSFLPRCYVILCRPE 1131
 QY 820 RNTPAYE 826
 Db 1132 LNTTEHF 1138
 RESULT 14

QY	131	YISVNVAVIGDNDSESWTVA	NFSLTELLRQITSAISDELDRKVRPALLETTPSADH	130
QY	131	YISVNVAVIGDNDSESWTVA <td>NFSLTELLRQITSAISDELDRKVRPALLETTPSADH</td> <td>130</td>	NFSLTELLRQITSAISDELDRKVRPALLETTPSADH	130
Db	58	YSPVLVAIGDPSDNRRAATTA	LLSPLVPIVSTIASSSETLSVARKQPSFLRTIPNDKYQ	117
QY	191	VEAVVQMLHFRNMWIIIVIV	SDPYGDNQGLGERVARDICIAEQETLP-TLPQNM	249
Db	118	VEVWVLLQKRGWMIWISLV	SGSDYDGLGVQALENQAQTGGICAFIDIMFSAQYGD--	175
QY	250	TSEERQRLVTVDKIQOSTAR	VVVYSSPDLTLHFENEVLRFNQTGAVMVIASSMAIDPV	309
Db	176	-----ERMOCMRLIAQAGAT	VVVVFSSROLARFFESVVLTLNIGCVWVASEMAISR	230
QY	310	LHNTELGHTLFGITGFIQSV	PIGFSDFRQMGA--GPPLSFTSQSYTCNOECDNC	366
Db	231	ITGPGIGIORIGMVLGVALI	QKRAVGLAFAFEARARADKEAPRCHKSMC-SSNQLCREC	289
QY	367	LNATLSFNTILRLSGERVY	SVSYAVYAVAHALSHLLSGDKSTCTKRKYVPMOLLEEIMK	426
Db	290	QAFMAHMPKRLAKSMSSAY	NAYRAYVAVHGLHQLLGCSAGCSRGRVPMOLLEQIHK	349
QY	427	VNFTLLDHOJFEFDQDGVAL	HLLEIVQNMQRSPQGSVAS--YPLQRLQNKNIQDISM	484
Db	350	VHFTLHKDTVAFNDNRDPL	SYNIIADMMGPKTFVYLSSSTWSPVLQNI-NETKIQMH	408
QY	485	TVNNITIPMSMCKSKQSGO	KKRPVGIHVCCFECIDCLPFTFLNTHTEDEYECQACPNNEWS	544
Db	409	GKDQVQPSVCSQDCLBQHQ	RVTMGRFHNCCECVPCQAGAFIUNKS-DLYRCQPGCKREMA	467
QY	545	YQSETSCPKROLVLEHMEAF	TIYVALIAALGFISTALIVTFRKHOTPIYVSAGSPMC	604
Db	468	PEGQOTCPRTVAFLEAREH	TSWILLANLTLTLGLTAGLFMHLDDTPVRSAGRLC	527
QY	605	FLMTFTLLVAWVVPVYGP	PKVSTCQKROLFPLCTICISCAVRSQIVCAFKAASR	664
Db	528	FLMGSLAAGSGSLXGFER	GPETPRACILRLQALFALGFTIFLISCLTVASPOLIIIFKSTK	587
QY	665	FPRAYSTWRYQGPYVSMAF	ITVLKMYIVYIGMLARQSHPRDDDPDKITIVSCNENYR	724
Db	588	VPTREHAMVQVNHGAGLFP	MSIAQQLICITLWLVVMTPLPAREYQRPBLVMECTET--	645
QY	725	NSL-----LENTSIDLL	LVSVGFSFPMGKELPTNNEAKFTILSMTFEFTSSVSLCTF	778
Db	646	NSLGFILAFLYNG-----	LLTSINAFCSYLGIDPEPNNENAKCYFSLIFNFVSMIAFTT	701
QY	779	MSAYSGVLTVTDLLVTVL	NLAISLIGFGCKCYMLFYPEBNTPAYFNSMIGOTYMR	836
Db	702	ASVYDQKYLPAANNMAGL	SLSSGFGFYLPKCYVILICRDLNSTEHFQASIDYTER	759
RESULT 15				
AAU00508				
ID	AAU00508 standard; Protein; 1059 AA.			
XX	AAU00508;			
XX	29-AUG-2001 (first entry)			
XX	Chicken calcium-sensitive receptor (CaR) protein.			
DE	Avian; chicken; calcium-sensing receptor; CaR; clone CID;			
XX	extracellular calcium homeostasis; parathyroid hormone; PTH;			
KM	serum calcium regulator; bone deposition.			
XX	Gallus sp.			
OS	Gallus sp.			
XX	Key			
FH	Location/Qualifiers			
FT	1..19			
FT	/label- Signal_peptide			
FT	1..611			
FT	/label- Extracellular_domain			
FT	/note- "Amino-terminal predominantly hydrophilic domain"			
FT	20..1059			
FT	/label- Mature_CaR_protein			
FT	Protein			

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:47:40 ; Search time 19.8343 seconds
(without alignments)
4061.679 Million cell updates/sec

Title: US-09-927-315-9
Perfect score: 4443

Sequence: 1 MGPRAKTCISFLFLWLAE.....ERNTPAYFNSMIGCYTMRD 838

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :
1: PIR_73:*
2: pir1:*
3: pir2:*
4: pir3:*
5: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1094	24.6	1078	2	A56715 calcium receptor (
2	1088.5	24.5	1079	2	A59362 calcium/polyvalent
3	1087.5	24.5	1085	2	SA0476 Ca(2+)-sensing rec
4	1079	24.3	1088	2	B56715 calcium receptor (
5	1012	22.8	858	2	JC7683 taste receptor T1R
6	724.5	16.3	915	2	A49874 metabotropic gluta
7	721.5	16.2	912	2	JH0563 metabotropic gluta
8	718	16.2	879	2	JH0562 metabotropic gluta
9	718	16.2	879	2	JC7160 metabotropic gluta
10	699.5	15.7	872	2	JH0561 metabotropic gluta
11	685	15.4	1171	2	A42916 metabotropic gluta
12	675.5	15.2	1180	2	JC2132 metabotropic gluta
13	675.5	15.2	1212	2	JC2131 metabotropic gluta
14	673	15.1	871	2	A46742 metabotropic gluta
15	672	15.1	908	2	I49142 metabotropic gluta
16	655.5	14.8	999	2	T77628 metabotropic gluta
17	654	14.7	1199	2	A41939 G protein-coupled
18	649	14.6	1218	2	S71376 glutamate receptor
19	632.5	14.2	1267	2	T21340 hypothetrical prote
20	423	9.5	551	2	T30806 metabotropic gluta
21	218.5	4.9	921	2	T51136 ionotropic gluta
22	218.5	4.9	923	2	F84732 ionotropic gluta
23	215	4.8	1039	2	T45779 probable ligand-ga
24	200.5	4.5	953	2	E84732 probable ligand-ga
25	200.5	4.5	1099	2	T16283 hypothetrical prote
26	186	4.2	950	2	T51134 ionotropic gluta
27	185.5	4.2	925	2	T06128 hypothetrical prote
28	185	4.1	975	2	A84550 probable ligand-ga
29	184	4.1	925	2	T51133 ligand gated chann

30	184	4.1	951	2	T51132 probable glutamate
31	177.5	4.0	1005	2	S33525 guanylate cyclase
32	176.5	4.0	960	2	JR0356 gamma-aminobutyric
33	176	4.0	933	2	C96495 probable ligand-ga
34	174.5	3.9	912	2	T51131 ligand-gated chann
35	169.5	3.8	941	2	T51135 ligand-gated chann
36	168.5	3.8	962	2	D86186 hypothetrical prote
37	160.5	3.6	1679	2	T15968 hypothetrical prote
38	158.5	3.6	976	2	T51137 ionotropic gluta
39	158	3.6	997	2	S33754 glutamate receptor
40	144	3.2	918	2	I58178 glutamate receptor
41	144	3.2	949	2	S19808 glutamate receptor
42	143	3.2	965	2	I51244 N-methyl-D-asparta
43	131.5	3.0	1125	1	ORURCP spectact receptor p
44	130.5	2.9	938	2	A46612 N-methyl-D-asparta
45	129.5	2.9	885	2	JN0335 N-methyl-D-asparta

ALIGNMENTS

RESULT 1

A56715
C:Species: Homo sapiens (man)
C:Date: 19-Oct-1995 #sequence revision 19-Oct-1995 #text_change 01-Dec-2000
C:Accession: A56715; S49341; A49419; B49419; C49419
R:Garrett, J.E.; Capuano, I.V.; Hammerland, L.G.; Hung, B.C.P.; Brown, E.M.; Hebert, J. Biol. Chem. 270, 12919-12925, 1995
A>Title: Molecular cloning and functional expression of human parathyroid calcium rec
A:Reference number: A56715; MUID:95279439; PMID:7759551
A:Accession: A56715
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1078 <GAR>
A:Cross-references: GB:U0759; NID:9683744; PIDN:AAA86503.1; PID:9683745
R:Pearce, S.H.S.; Thakker, R.V.
A:Reference number: A56715; MUID:95279439; PMID:7759551
A:Accession: A56715
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-180, 'Q', 182-989, 'R', 991-1078 <PEA>
A:Cross-references: EMBL:X81086
R:Pollak, M.R.; Brown, E.M.; Chou, Y.H.; Hebert, S.C.; Marx, S.J.; Stelmann, B.; Lev Cell 75, 1297-1303, 1993
A>Title: Mutations in the human Ca(2+)-sensing receptor gene cause familial hypocalci
A:Reference number: A49419; MUID:94094324; PMID:7916660
A:Accession: A49419
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 178-192 <POL>
A:Experimental source: family N
A>Note: sequence inconsistent with nucleotide translation
A>Note: sequence modified after extraction from NCBI backbone
A>Note: 186-Arg mutation is associated with familial hypocalcemic hypercalcemia and
A>Note: sequence extracted from NCBI backbone (NCBIN:142453)
A:Accession: B49419
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 289-303 <PO2>
A:Experimental source: family E
A>Note: sequence modified after extraction from NCBI backbone
A>Note: 298-Lys mutation is associated with familial hypocalcemic hypercalcemia and
A>Note: sequence extracted from NCBI backbone (NCBIN:142455)
A:Accession: C49419
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 788-802 <PO3>
A:Experimental source: family J
A>Note: sequence modified after extraction from NCBI backbone
A>Note: 796-Tyr mutation is associated with familial hypocalcemic hypercalcemia and
A>Note: sequence extracted from NCBI backbone (NCBIN:142457)

C:Keywords: glycoprotein; receptor; transmembrane protein

Query Match 24.6%; Score 1094; DB 2; Length 1078;

Best Local Similarity 31.5%; Pred. No. 1.6e-75; Matches 279; Conservative 160; Mismatches 342; Indels 106; Gaps 26;

```

QY 16 WIAEPAENSDFYLP-----GDYLLGLGFLSHANKGIYHNLQVPMCKEYEVIG 68
D 8 WVLLATLWHSAYGPDORAKKGDIIIGLGPPIHF---GVAAKD--QDLKSRPSEVICR 62
QY 69 YNL-----MAMRAVEEINNDSSLPGVLLGLEYIVDC-YISNNQPVLYFLA-HEDNL 121
D 63 YNBERFWLQAMTAIEEINSSPALLPVLTYGRIPTCNTVSRALEXTLSFVAKNNIDS 122
QY 122 LPIDQ--DYSNYISRYVAVIGPDNSESWYVAFNLSLELPQITYSASIDELRDKVFP 179
D 123 LNLDFPCNSEHISTIAVVGATGSGYSTANANLGLFYIQVGYASSRLSKNPKFS 182
QY 180 LKRTTSPADHVEAMVOLMELFRNNWIIIVSSPTYGRDNCQLIGERYARDICIAFOET 239
D 183 FLRTIPNDEHQATAMADIEEFRRNNWVTIADDDYGPRIEKREREAEREDICIDFSEL 242
QY 240 LPTIQPNQNTSEEROLRYTVDKLOOSTAVYVYFSDLLHYEFNEVLRQNTGAWMI 299
D 243 I-----SOYSDDEEIQHVEYI---QNSTAVYVYFSSGPDLEPLIKELRYRNTGRIWL 294
QY 300 ASESMAIDPVHLNLTGLHGTFLGITIQSVPIGPSEF-----REMGPO 344
D 295 ASEMASSSLIAMPQYFHVVGITGPALAKAGIPIGFRFLKVKHPRKSVHNGFAKEMEE 354
QY 345 A-----GPPPLRTSOSTYCNQECNCLNATSFNTILSGE-----382
D 355 TFNCHLOEGANGPLPVDTFLNGH--EESGDRFSSNSTAFRLP--CTGEMNISVETPEYID 410
QY 383 -----RVVYSYSAVYAAVAHALSHLGC-----DKSTCKRKYVPMQLEEIMKYN 430
D 411 YTHIRISYNYLAVYSTAHALQDIYTCLPGNGLFTNSCADIKVEAMQVLAKHRLHNF 470
QY 431 -LIDHQIFPDQDVALHLEIVOMQDRSONP--FQSVASYPL-----OROLKNIDIS 482
D 471 NNMGEQVTFDECGDLVGNYSIIINHLSPEDGSIVKEY-GYVNYAKKGERLFINKEKIL 529
QY 483 WHYNNNTIPSMCKRCOSQKKRPV-GIHYCCFECIDCLPGTLNTEDEYECQACPN 541
D 530 WSGSREVPSCNRCDLAGRKGIIEGEPCECEVCPEGEYSDEF--DASACNKCPDD 588
QY 542 EMSYQSESCFKROLVEFHEHAPLIAVALLAALGPLSTAILIYFNHPTPIVRSAG 601
D 589 FMSNENTSCIAKEIEFLSTPEFGIALITLFAVIGITLAVLVLFKFRNPTLYAKNR 648
QY 602 PMCFMLTLVLVAVYVYVVPKPVSTCLQALFPLCFCTICISCIASVRSFOIVCAF-- 659
D 649 ELAYLLFLSLICFSSSLFIFGERPDWTCRROPAGISFVLCISCLIKYNNRLVAFEA 708
QY 660 KMASRFRPANSYWRVYGPYSMAFITVLAQVYVIGMLARPOSHPTDPDKITVSC 719
D 709 KIPISFHRK--MMGLNQLFLVLCFEMQIVICVIMLYTPAPSSYNOELEDIITYC 765
QY 720 NPNRNSILFNTSLDILVYVGFPAWNGKELPTNYNEAKRITLSTMFYFSSLSCTEM 779
D 766 HEGSLMALGFLIGTCLLAACFFAFKRSKRLPENENAKRITLSMLIFLWIS---FI 822
QY 780 SAVSGVLTVYDILLVYVNLALSLG---YFGPKCYMILEYPERNT 822
D 823 PAYASTGKRFVS-AVEVIALAASGELLACIFENKITYITLEKPSRNT 868

```

RESULT 2

159362

calcium/polyvalent cation-sensing receptor precursor - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 05-Nov-1999

C:Accession: I59362; A55594

R;Ruut, M.; Molliver, M.E.; Snowman, A.M.; Snyder, S.H.
Proc. Natl. Acad. Sci. U.S.A. 92, 3161-3165, 1995

A:Title: Calcium sensing receptor: molecular cloning in rat and localization to nerve

A:Reference number: I59362; KUID:95241465; PMID:7724534

A:Accession: I59362

A>Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1079 <RES>

A:Cross-references: EMBL:U02829; NID:g790578; PID:AC52195.1; PID:g790579

A:Experimental source: striatal

R;Riccardi, D.; Park, J.; Lee, W.S.; Gamba, G.; Brown, E.M.; Hebert, S.C.
Proc. Natl. Acad. Sci. U.S.A. 92, 131-135, 1995

A:Title: Cloning and functional expression of a rat kidney extracellular calcium/poly

A:Reference number: A55594; KUID:95116508; PMID:7816802

A:Accession: A55594

A:Molecule type: mRNA

A:Residues: 1-133, 'X', 135-1079 <RIC>

A:Cross-references: GB:U010354

A:Experimental source: kidney

C:Keywords: calcium; glycoprotein; phosphoprotein; transmembrane protein

F;1-20/Domain: signal sequence #status predicted <Sig>

F;187-212/Region: hydrophobic

F;613-635/Domain: transmembrane #status predicted <TM1>

F;630-670/Domain: transmembrane #status predicted <TM2>

F;683-700/Domain: transmembrane #status predicted <TM3>

F;725-744/Domain: transmembrane #status predicted <TM4>

F;770-790/Domain: transmembrane #status predicted <TM5>

F;806-828/Domain: transmembrane #status predicted <TM6>

F;90,261,287,386,468,488,594,893,1005/Binding site: carbohydrate (Asn) (covalent) #st

F;784/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status pred

F;899,901/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status pred

Query Match 24.5%; Score 1088.5; DB 2; Length 1079;

Best Local Similarity 31.7%; Pred. No. 4.3e-75; Matches 284; Conservative 164; Mismatches 328; Indels 121; Gaps 31;

```

QY 10 SLFLLVTLAEPAENSDFYLP-----GDYLLGLGFLSHANKGIYHNLQVPMCKEY 62
D 9 ALLALAW-----HSSAYGPDQRAKKGDIIIGLGPPIHF---GVAAKD--QDLKSRPE 56
QY 63 EKVYIGYNL-----MAMRAVEEINNDSSLPGVLLGLEYIVDC-YISNNQPVLYFLA 116
D 57 SVECIIRYNFGEFWLQAMTAIEEINSSPALLPVLTYGRIPTCNTVSRALEXTLSFVA 116
QY 117 -HEDNLPIQD--DYSNYISRYVAVIGPDNSESWYVAFNLSLELPQITYSASIDELRD 173
D 117 QNKIDSLNDEFCNSEHISTIAVVGATGSGYSTANANLGLFYIQVGYASSRLSLN 176
QY 174 KVRPALLRTTSPADHVEAMVOLMELFRNNWIIIVSSPTYGRDNCQLIGERYARDIC 233
D 177 KNQKSLRLTIPNDEHQATAMADIEEFRRNNWVTIADDDYGPRIEKREREAEREDIC 236
QY 234 IAPQETPLTQPNQNTSEEROLRYTVDKLOOSTAVYVYFSDLLHYEFNEVLRQNT 293
D 237 IDESELI-----SOYSDDEEIQHVEYI---QNSTAVYVYFSSGPDLEPLIKELRYRNT 288
QY 294 TGAVWISSEMAIDPVHLNLTGLH--LGTFLGITIQSVPIGPSEF-----338
D 289 TGRITLASEMASSSLI-AMPEYFHVVGITIGRLAKGQIPGREFLQKHPKRSVNGF 347
QY 339 -REMGPOA-----GPPPLRTSOSTYCNQECNCLNATSFNTILSGE-----382
D 348 AKFWEETFNCHLOEGAKGGLPVDTFVRSH--EEGGRNLLNSSTAFRLP--CTGEMNIS 403
QY 383 -----RVVYSYSAVYAAVAHALSHLGC-----DKSTCKRKYVPMQLEE 423
D 404 VERPYMDYEHRLRISYNYLAVYSTAHALQDIYTCLPGNGLFTNSCADIKVEAMQVLA 463
QY 424 IMKYNFT-LIDHQIFPDQDVALHLEIVOMQDRSONP--FQSVASYPL-----OROL 475
D 464 LRHLNFTNNMGEQVTFDECGDLVGNYSIIINHLSPEDGSIVKEY-GYVNYAKKGERLF 522

```



```

OY 476 KRIDISMHVTNNTIPMCMSCROOSGQKKRY-GIHVCCFFCICDCLBGTFLNHEDEYE 534
Db 533 INEEKIIIMSGFSREYFENCSDRCOAGTRKGIIEBPCCFCBECBPGEYSGEF-DASA 581
OY 535 COACNNMESVOSSEKCRKRLQVLEWHEBAPITAAVLLAAGFESTLAILVTFMRHPQR 594
Db 562 CDKCDDDWMSNNHNSHCIAKETEFIAWTEPEGIALTLTFLAVLIGITFLAVLVGFIERNRP 641
OY 595 IVRSAGBMCFLMTLLILVAVYMWVYVYGPSPYCLCRLQALFPLCTFICICIASRSHQ 654
Db 642 IYKANNRREIYLLLESLOCFSSESLFEIGEPQDMWCRRLQRPAGFISFVLCICIIYKTR 701
OY 655 IYCAF-KHASFEPRAYSIWRYOGPYVSMATIVLKVIVYIGMLARPQSHPRDPDP 712
Db 702 VLVLEAKIPTFSHRK---WMGLNQFLVFLCTEPMLICIIIMLYTAPPSRYNHELD 758
OY 713 KTIIVSCBNPNYNSLFTNTSL---DLILSVGFSFAYMCKELPTYNNEAKFTTILSKTFY 769
Db 759 ELIIFTTC---HGSISMAJSLIGYTCLLAAICOFPAFRSKRKBENFNEAKFTTFMLTFE 815
OY 770 TTSVEICTFMSAYSGVLVTYIDLVTYVNLILASIG----YBPRCYMILPEBENT 822
Db 816 IYWIS---PIRAYASTYGRFVS-AVEVAILLASGLLACIFENKVVYIILEPSSNT 868

```

```

RESULT 3
S40476
Ca(2+)-sensing receptor - bovine
C:Species: Bos primigenius taurus (cattle)
C:Date: 19-Mar-1997 #sequence_revision 19-Mar-1997 #text_change 05-Nov-1999
C:Accession: S40476
R:Brown, E.M.; Gamba, G.; Riccardi, D.; Lombardi, M.; Butters, R.; Kilfor, O.; Sun, A.;
Nature 366, 575-580, 1993
A:Title: Cloning and characterization of an extracellular Ca(2+)-sensing receptor from
A:Reference number: S40476; MUID:94077182; PMID:8255296
A:Accession: S40476
A:Status: Preliminary
A:Molecule type: mRNA
A:Residues: 1-1085 <BRO>
A:Cross-references: GB:S67307; NID:g453108; PID:AMB29171.1; PID:g453109

Query Match      24.5%, Score 1087.5; DB 2; Length 1085;
Best Local Similarity 31.6%, Pred. No. 5.2e-75;
Matches 279; Conservatve 161; Mismatches 347; Indels 97; Gaps 28;

OY 11 LEFLMVALEPRAENDFLPYGDYLLGGLFSLHANKGIYHLNFIQV--MCKEYEVKYI 67
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 11 LAFSTWCISAGPDDRAOKKGDIIIGGLPIHFGV-AKQDDLKRPSEVECIKRNFR-- 67
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 68 GYNLMQAMRFAVEEITINDSLIPVLLGEIYDVC-YISNNYPVLYFLA-HEDNILPIQ 125
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 68 GFRMQAMIFALEEINSSPALPNMTLTGRIFDTCNFSKALEALISFAQKRIISLND 127
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 126 E-DYSNTISRVAVIGDENSEVMTVANFLSTFLPIITYSAISDELBDKVRPALLRT 183
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 128 EPCNCSEHPISTIAVVGATGSGISTAVANLLGLFTIPOVSTASSRSLSNKQKFSFLT 187
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 184 TPSADHVEAVYQLMLHFRMNNIIVLVSDTYGRNGQLGRVARRDICLFOETPLTL 243
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 188 IPNDHQATAMADIIEFRMNNVGTIAADDVGRGIEKREPREAERDICIDEFSELI--- 244
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 244 QPNOMTSEERORLTIYDKLQOSTARVYVPSPLTLXHFNEVLRQNTFAVNIASGS 303
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 245 ----SQYDDE-EKIDQVVEVIONSTAKVIIVSSGPDLEPLIKETIVRNITGRIMLASBA 299
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 304 WAIDPVNLNLTELGH-LGTFLLGTITQSVPIPGFSEF-----REMGPQA-- 345
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 300 WASSSLI-LAMPFYHVGVTIGFGKAGQIPGRFELGLKVHPRKSVHNGFAEFMEETFN 358
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
OY 346 -----GPPPL-----SRTSOSTY----CNOECDNCLMTLSFTNITLRSG 381
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 359 CHLGAGAGPRLPVDFELRGHEGGARLSNPSPTAFRPLTGE-ENISVETTPYMDTHTL-- 415
   | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

```

```

0Y 362 ERVYVSYSANVAVNAHLSLGC-----DKSCTKRVYPMQMLEIKKAVFTL-L 432
Db 416 -RISNYNLAAYSIAHMLQDLYTTCIPGRGLFTGSGOADIKKVAMOVLAHHLNFTSNM 474
0Y 433 DHOJFDPQGVVAHLLEIVQOMQDRQNP--FOSVASYPL-----OROLKNIDQISWHT 485
Db 415 GEQYTFDECGDLAAGSIINHLSPEDGSIYFEV- GYIVAVAKKERLFINDKILMSG 533
0Y 486 VNNTIPMSCKSRQOSQOKKPV-GIHVCCFECIDLPTGFLNTHDEYECQACPNEMS 544
Db 534 FSRREVPFSCSRDCLAGTKRGIIEGEPTCCFEVECPDEGYSDET-DASACDKCPDDMS 592
0Y 545 YOSESCRRQVLELVENHKAPTAVALALAGLSTALIVEMHROPIYRSAGPWC 604
Db 593 NENHTSCAKIEFLMSTPEPGIALTFAYALGFLVAVYIKRNTPIYKATNRELS 652
0Y 605 FLMLTLLVAAYVVPVYVGPKVSTCLRQALEPLCTTCIOSCIAVRSQIYCAF-KMA 662
Db 653 YLLFSLSLCFFSSSLFEGEPQDCTRLPAPGISPVLCISCIILTKNRLVFEAKPI 712
0Y 663 SRFPRAVYWRVYQGPVYSAFTVLKMYIVYIGMLARQSHRDRPDPKRTIYSCNP 722
Db 713 TSPHRK---WMLGMLQFLVFLCTPEMIOICAIWLTAPSSYRNHLEDELITFTCHG 769
0Y 723 YRNSLHFTSLDLLLVSGSFAYAGKELPTYNENAKFTILSTMFTFYSVSLCTFMSAY 782
Db 770 SLMLAGFLIGYTCLLAICFPFAKRSKRLPENENAKFTILFSLILFFIYWIS---FIRAY 826
0Y 783 SGVLTYVDLLVYVNLALSLG---YFGPKCYMLTFPERNT 822
Db 827 ASYGRKEVFS-AVEVIALIASFGILACIFNKNRYIILFYSRNT 869

```

[illegible]

```

0Y      24  NSDSELDGDTLGGELFSLHANMKKIVHNLQFV---MCKEYEVKIVGNILMQAMRAVE  80
          :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
          :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db      26  SQQKAGSDYLLGGELFPFGSTEET--LNORTQPNISPCNRF--SPGLFLPAMAKMAVE  81
          :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
0Y      81  EINDSSILGCVLLGYEIVDVCYISNNVQPY-----LYFLAH-EDNLLPLQEDYSWYIS  133
          ||||| :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

```

RESULT 6
A49874
metabotropic glutamate receptor 7 - rat
N:Alternate names: metabotropic glutamate receptor mglur7
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Jul-1996 #sequence.revision 02-Jul-1996 #text.change 24-Sep-1999
C:Accession: A49874; 157954
R:Okamoto, N.; Horii, S.; Akazawa, C.; Hayashi, Y.; Shigemoto, R.; Mizuno, N.; Nakanishi, S. *Biol. Chem.* 269, 1231-1236, 1994
A:Title: Molecular characterization of a new metabotropic glutamate receptor mglur7 c
A:Reference number: A49874; MUID:94117433; PMID:8288585
A:Accession: A49874
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1,915 <RMS>
A:Cross-references: GB:D16817; NID:g458728; PIDN:BAA04092.1; PID:g458729
R:Saugstad, J.A.; Kinzle, J.M.; Mulvihill, E.R.; Segerson, T.P.; Westbrook, G.L.
Mol. Pharmacol. 45, 367-372, 1994

A:Title: Cloning and expression of a new member of the L-2-amino-4-phosphonobutyric acid
 A:Reference number: 157954; MID:94195260; PMID:8145723
 A:Accession: 157954
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-915 <RES2>
 A:Cross-references: EMBL:006832; NID:g459657; PIDN:AAA20655.1; PID:g459658
 C:Genetics:
 A:Gene: MGLUR7
 C:Superfamily: metabotropic glutamate receptor 4
 C:Keywords: neurotransmitter receptor

Query Match 16.3%; Score 724.5; DB 2; Length 915;
 Best local similarity 28.0%; Pred. No. 3e-47;
 Matches 257; Conservative 152; Mismatches 353; Indels 157; Gaps 41;

```

QY 9 CSLEFLWVLAEPNSDFLP-----GDYLLGLFSLHANKGIIVHLNLTQVPMCKEY 62
DB 19 CVLEVLACVLAARQGEVAPHSIRIEGDVTLGLFVPAH--KGP-----SGVP-CGDI 70
QY 63 EVKVGYNLMQAMRFAVEEINDDSLPGVLLGYEIVDVCYISNNQVPLYFLAHEDNLL 122
DB 71 K-RENGIHLEALYALDDINDPNLLPVTLGARLIDTC--SRDTYALQSLTFYQAL- 126
QY 123 PLOEDVSNY-----ISRYVAVIGPDNSSESVMTVANFLSLFLLPQITYSASIDE 170
DB 127 -IQKDSVRCINGEPPEVVEKPEKVGYGAGSSYSIVANILRFQIPQISTASTABE 185
QY 171 LNDKVFAPLLPTPSADHVEAMVQLMFRMNIIVLVSDTYGRDNGQLGERVAR 230
DB 186 LSDRRYEDFSRVPPDSFOAQMVDIVKALGMNVSTLASBGSV-EGVSEFQISKE 244
QY 231 --DICIAPETLPTQPNQMTSEERQV----TIVDKLOOSTARVAVVPSBDLTLYNF 284
DB 245 AGGLCTAQSVRIP-----QERKRDITIDRIKQLDLPNSAVVIFANDEDIKOI 295
QY 285 FNEVLNQNTG-AVMIASESMA--IDPVHLNLTGLTFLGITQ--SVPIGFSER 339
DB 296 LAAKADQVGHFLWVSGDSKINP-LHQHEDIAE---GATIQPRARVGEADAF 350
QY 340 EHGPPQAPPLPSTOSYTCN-----QECNDC--LNAITSEFTILSG-ERY---- 384
DB 351 -----TSRTLENNRNWFAEYWEENFCKLTIGSKREKEDRCKTQGERIGKDS 400
QY 385 -----VYSVSAVYAVAHLSLLGCDKSTCK-RVYYP-----MOLLSEIKVNF 429
DB 401 NTEQESKQVQVLDAYATAMAHALHHM--NKDLCADRYGCEPEQAGGKLLKTYIHVNF 457
QY 430 T-LLDHQIFEDPOGVALLHLEIVQWQDRSQNPQSVASYPLQROLKNIODISMHTVNN 488
DB 458 NSAGTPVFNENKNGDAPGRYDIFQYQTNTNPGYRLIGQWTDQL- NIDEMQMGKVR 516
QY 489 TTPMSKCRQSGQKKKRVGIIHCCFECIDCLPGTFNLHTEDEVEQACPNNEKSYOSE 548
DB 517 EIPSSVCTLPCKPGQKKKQKGTGCCWCEPC---DGYOYQDEMCQCHCPYDQRENER 573
QY 549 TSCFQKQVLEFMHEAPTAVALLAALGFLSTLALIVIRHFORPIVASAGPMOFLML 608
DB 574 TGCQNPITIKLEHSHMAVIVFLALGLITATLEWATIRINDPIVASGELESTVLL 633
QY 609 TLLAVAMVVPYVGPVKSTCLRQALPLCTFCISCIAY-----RSFQVYCAFKMAS 663
DB 634 TGLPILCYITITFLIAKPDVAVCSFRVFLGL--GMCISYALALLTKNRIRYRFEQKKSV 691
QY 664 RRPRAVSVKRYQGPVSNAFITV-LKMYIVVIGMLARPOSHPRDDPKTYI-----V 717
DB 692 TAPRLISPTSLA---ITSLISVOLGLVIFWGV-----DDPNIIIDDEKH 736
QY 718 SCNPYRNSLLENTSLDL-LLSVGS-----FAVNGKELPTNVEAKFTILSMTF 767
DB 737 TAMPBQARVYLCDDITDILQSLGYSILLAMVTCYVYAKTKGVPENNEAKPIGFTM-- 794
QY 768 YFTSSVSLCTFMSAVSGVLVTLIVDL-----LVTVNLILA-ISLG-YFQPKCYMILYPE 819

```

```

DB 795 -YTTCTVMALEPIPIFGTRQSAEKLYITQTYTLTISMNLSASVALGMLVPMKYIIIFHPE 853
QY 820 RMT---PAYFNSMIOGYTM 835
DB 854 LNVQKRKRKRSFKAVVYAAATM 872

```

RESULT 7
 JH0563
 metabotropic glutamate receptor 4 precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
 C:Accession: JH0563; 158149
 R:Tanabe, Y.; Masu, M.; Ishii, T.; Shigemoto, R.; Nakanishi, S.
 Neuron 8, 169-179, 1992
 A:Title: A family of metabotropic glutamate receptors.
 A:Reference number: JH0563; MID:92110002; PMID:1309649
 A:Accession: JH0563
 A:Molecule type: mRNA
 A:Residues: 1-912 <TRAN>
 A:Experimental source: brain
 R:O'Hara, P.J.; Sheppard, P.O.; Thøgersen, H.; Venezia, D.; Haldeman, B.A.; McGrane, Neuron 11, 41-52, 1993
 A:Title: The ligand-binding domain in metabotropic glutamate receptors is related to A:Reference number: 158149; MID:93332699; PMID:8338667
 A:Accession: 158149
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-123, 'R', 125-912 <RES>
 A:Cross-references: GB:M00518; NID:g205400; PIDN:AAA93190.1; PID:g205401
 C:Comment: This protein is coupled to a G protein and evokes a variety of functions b
 C:Genetics:
 A:Gene: GLUR4
 C:Superfamily: metabotropic glutamate receptor 4
 C:Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; phos
 F:1-32/Domain: signal sequence #status predicted <SIG>
 F:33-912/Product: metabotropic glutamate receptor 4 #status predicted <MET>
 F:588-610/Domain: transmembrane #status predicted <TRI>
 F:625-645/Domain: transmembrane #status predicted <TRI>
 F:657-675/Domain: transmembrane #status predicted <TRI>
 F:700-720/Domain: transmembrane #status predicted <TRI>
 F:751-772/Domain: transmembrane #status predicted <TRI>
 F:786-807/Domain: transmembrane #status predicted <TRI>
 F:822-841/Domain: transmembrane #status predicted <TRI>
 F:98,301,454,484,569/Binding site: carboxylate (Asn) (covalent) #status predicted
 F:621,689,695,859,870/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 16.2%; Score 721.5; DB 2; Length 912;
 Best local similarity 26.5%; Pred. No. 5e-47;
 Matches 240; Conservative 161; Mismatches 349; Indels 155; Gaps 37;

```

QY 8 ICSLEFLWVLAEPNSDFLP-----GDYLLGLFSLHANKGIIVHLNLTQVPMCKEY 59
DB 16 ICSLEFLWVLAEPNSDFLP-----GDYLLGLFSLHANKGIIVHLNLTQVPMCKEY 67
QY 60 KEVEKVGYNLMQAMRFAVEEINDDSLPGVLLGYEIVDVCYISNNQVPLYFLAHED 119
DB 68 GEIK-KEKGIHLEALYALDDINDPNLLPVTLGARLIDTC--SRDTYALQSLTFYQAL 124
QY 120 NLPLOEDVSNY-----ISRYVAVIGPDNSSESVMTVANFLSLFLLPQITYSASIDE 167
DB 125 AL--TEKDSVRCINGEPPEVVEKPEKVGYGAGSSYSIVANILRFQIPQISTASTABE 182
QY 168 SDELBDKVFAPLLPTPSADHVEAMVQLMFRMNIIVLVSDTYG----- 216
DB 183 APDLSDNSRYEDFSRVVPSDQAOAMVDIVKALGMNVSTLASBGSVGEVFAFIQKS 242
QY 217 RONGQLGERVARARQICIAFOETLPTQPNQMTSEERQVLYTDKLOOSTARVAVVPS 276
DB 243 RENG-----GVCIASQSVKIRP-EKPTGEFPIKIRL-----LETNARKGIIIFVA 285
QY 277 PDLTYHFPEFNEVLNQNTG-AVMIASESMAID--PVLHNLTEGLTFLGITYI--QSV 331

```

Db 286 NEDDIRVLEAARANDGHEFFMGSDWSGSKAPVL-RLEEVAE---GAVTILPRKMS 340
 QY 332 IPGSEFREMPOAGPPLISRT-----SOSYCNOCDCDNLATSFNTILRLSG 381
 Db 341 VRGDRIFS-----SSTLDNNRRNRTWAEFEMEDNEFKLSRRLAKGSHIKCTN 390
 QY 382 -ERV-----VYSVSAVAVAHLSLGLCDKSTCKRY-----YYPMLL 421
 Db 391 RERIGQDSAYEQEGKVQFVIDAVYAMGHALHAM---HRDLPGRVGLCPRMDPVDGQLL 447
 QY 422 EELKRVFT-LLDHQIFPDGVALHLELYQOMDRSONPQSVASTYP-LQOLKNITQ 479
 Db 448 KYIRNVFSGIAGNPVTFENMGDAPGRYDIYQOLRNGSARVYIGSTWDLHLRIERMQ 507
 QY 480 DISHNTVNTIPMSGSKRCSGOKKRPVGIHVCCFECIDCLPGFNLHTEDEYEQAC 539
 Db 508 ---WPGSQQLPRISCLPCQPERKTYKGMACCWCEPC---TGIOYQDRYTKCTCP 561
 QY 540 NNEWSYSETSCFRQQLVLEMHAPPTAVALLAFLSLTALVTFWRHPTPIVISA 599
 Db 562 YDMPTERTSCOPRIPIKLEMDSPMAVLPFLAVGIAATLFFVVFERYNDPIYKAS 621
 QY 600 GGPICFMTLLVAVYVPPYVGPVKYSCQLCQALFPLCTICISCIANRSQIYCAF 659
 Db 622 GRELSYLLAGIFLCYATTFMLIAEPDLGTCSLRIFLGCMSTSYALLTKTYRIRIF 681
 QY 660 KMASRFPRAVSYYWRYQSP--YVSMAFITV-LKAVIVYIGMARPOSHPTDPPDKITI 716
 Db 682 EQGRKVSAA---PRFISPAQQLAITFLLISLQILGICVWVNDP-SHSVVDPPDOR--- 733
 QY 717 VSCPNRNSLLENTSIDL-LLSVGFS-----FAYGKELPTNYNAKFTTYSMT 766
 Db 734 -TIDPRFARGVLCDDISLILCLGLSMLMTCTVYAITRGVPEFNFNAKIGFTM- 791
 QY 767 FYFTSSVSLCTFMSAYSG-----VLVTYDVLVTVNLALISLG-YFGKCMILFEP 818
 Db 792 --YTCTIWMIAFIFIFRGTSGSADKLIYQTITLVSVLSASVSLGMLYMKYITILFHP 849
 QY 819 ERNRP 823
 Db 850 EQNVP 854
 RESULT 8
 JH0562
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 13-Sep-1998
 C:Accession: JH0562
 R:tanabe, Y.; Masu, M.; Ishii, T.; Shigemoto, R.; Nakanishi, S.
 Neuron 8, 169-179, 1992
 A:Title: A family of metabotropic glutamate receptors.
 A:Reference number: JH0561; MUID:92110002; PMID:1309649
 A:Accession: JH0562
 A:Molecule type: mRNA
 A:Residues: 1-879 <TRAN>
 A:Experimental source: brain
 C:Comment: This protein is coupled to a G protein and evokes a variety of functions by m
 C:Superfamily: metabotropic glutamate receptor 4
 C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane prot
 F:1-22/Domain: signal sequence #status predicted <Sig>
 F:23-879/Product: metabotropic glutamate receptor 3 #status predicted <MET>
 F:377-399/Domain: transmembrane #status predicted <TRI>
 F:614-634/Domain: transmembrane #status predicted <TRI>
 F:646-664/Domain: transmembrane #status predicted <TRI>
 F:689-709/Domain: transmembrane #status predicted <TRI>
 F:735-756/Domain: transmembrane #status predicted <TRI>
 F:770-791/Domain: transmembrane #status predicted <TRI>
 F:804-828/Domain: transmembrane #status predicted <TRI>
 F:209-292,414,439/Binding site: carbohydrate (Asn) (covalent) #status predicted
 F:610,845/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 16.2%; Score 718; DB 2; Length 879;
 Best Local Similarity 26.1%; Pred. No. 8.9e-47;
 Matches 225; Conservative 147; Mismatches 357; Indels 134; Gaps 29;
 QY 29 LPQYLLGLGFLSHANKGIYHLNLFQVPMCKEYEVVIGYNLMQMRFAVEELINDSL 88
 Db 35 IEGDLVGLGFLPIENKGTGEEGRINEDR-----GIORLEAMLFADENKNDYL 85
 QY 89 LPGLVLEYEIVDC-----YISNVQPVLYFLAHEDN---LLP-----IOEDYSNTISRY 135
 Db 86 LPQVGLVHLIDTCSRTYALBQSLFVRSALTFVDAEAEWCPDQSAIDE---NIPILI 142
 QY 136 VAVIGPDNSESVMYANFLSLFLPQITYSASIDELDKYRFPALLFTPSADHVEAMV 195
 Db 143 AGVIGSGSVSYSLQVANILRFQIPQISYASTSAKLSDKSRDYVFARTYPPDFQAKAMA 202
 QY 196 QLMHFPMNNTIYLVSSDTYGRDNGQLGERVARNQICAPQELPILQONMNTSPERQ 255
 Db 203 EILNFFMWTYVTVASGDIGETGIEAFEDQEARLRNICIATAEKVGSRNIRKSYDS--- 258
 QY 256 RLVTYVDKLOQSPARVVVF---SPDLTYHFENEVLRONTGAVWIASMAIDPVLHN 312
 Db 259 ---VIRELQKPNARVVVLFMRSDSRELIANAARV-NASFT---WVASGQWQAGESIVK 311
 QY 313 LTELGHILGTLGITIOSVPI-----PGSEFREMGPQAGPPLISRTS 354
 Db 312 GSEHVAAGA-ITTELASHPQFDRYQSLNPNYNNRHPFRFWE----- 356
 QY 355 QSYTCN-----QEDCNCLNLTSLFNTILRLSGEVRVYVSVAVAVAHALHSLG-- 404
 Db 357 QKFCSLQKNRNRQVCDKHLAIDS-----NYBQESKIFVYNAVYAMAHALHKMORTL 411
 QY 405 CDKST--CTRVVYPMQL--EELIKVNTL-----LDHQIFPDGVALHLELYQW 453
 Db 412 CPNTTKLDAMKILDGKKLYEYLLKINFAPENPKAGDAIVKFDFFGDMGRVNFNL 471
 QY 454 QMRSQNPQSVASYIYLOLQKNIDISMTHTVNTIPMSGSKRCSGOKKRPVGIHV 513
 Db 472 QQTGKYSYLRKGM--AEFLSLDVDSIHMS--RNSVPTQCSDDPCAPNEMKKNQPGDVC 527
 QY 514 CFECDIDLPGLNHTDEDECQACPNNEWSYSETSCFRQQLVLEMHAPPTAVALLA 573
 Db 528 CWITICPEPEYEL---VDETCMDCGPGQMPADLSCLNLPEDYIMEDMAAIGPTTIA 584
 QY 574 ALGFLSTLAILVFMHFGQPIVRSAGPMPCTMLTLLVAVYVPPYVGPVKYSTCLR 633
 Db 585 CLGFLCICYITVYIKRNNPPLVYASGRELCYILFVSVLSYCTFFFIAPKSPVICALR 644
 QY 634 QALFPLCTICISCIANRSQIYCAF---KMASRFPRAVSYYWRYQSP---YVSMAFIT 666
 Db 645 RLGLGTSFAICYSALLTKTKCIARIFDGVKNGAQRPFIS-----PSSQVFLCLGLL 697
 QY 687 VLKAVIVYIGMARPOSHPTDPPDKITTVSCPNRNSLLENTSIDLISLVGSEFAY 746
 Db 698 VOIYVMSVWMLLETPGRRTTLPEKRETYILKCNVK--DSGMLSLITVDDVLYILCYTAF 756
 QY 747 MGKELPTNYNDEAKFTTISMTFTTSVSLCTFMSAYSGVLTIVDLYL-VLNLALISLG 805
 Db 757 KTRKCPENFENAKFIGFTM---YTCTIWMIAFIFIF---YVTSDDYVQVTTMCSIVSL 810
 QY 806 -----YFGKCMILFEPERN 821
 Db 811 GFVVLGLFLAPKVAITLFPQKN 833
 RESULT 9
 JC7160
 C:Species: Mus musculus (house mouse)
 C:Date: 04-Mar-2000 #sequence_revision 04-Mar-2000 #text_change 11-May-2000
 C:Accession: JC7160
 R:Minoshima, T.; Nakanishi, S.
 J. Biochem. 126, 889-896, 1999

```

0Y 29 LRGDVLGGGLSLHANNMKGIHNLNFQVPMCKEYEVKTGYNLMQAMFAVEIINDDSL 88
Db 35 IEGDVLGGGLSLHANNMKGIHNLNFQVPMCKEYEVKTGYNLMQAMFAVEIINDDSL 85
0Y 89 LRGVLLGXEIYDVC-----YISNNVQPLVFLAHEDN---LLP-----IQEDSYNISRV 135
Db 86 LRGVLLGHIIDTCGRDLYALGSLFEPYRASLTKYDEAEYMCPOGSAIQE---NIPLLI 142
0Y 136 VAVIGPDNSSEVMYANFLSLFLPLQITYSAISDELDRKVRFPALLRTTPSADHHEVM 195
Db 143 AGVIGGSSVSIVQANLRLFLQIPQISASTASAKLSKRSYDFARVPDPDYQAKAMA 202
0Y 196 OLMFERNMIIIVYSSDTYGRDNQGLGGEVARDIDICAOETPLQTPQNMNTSEERQ 255
Db 203 ELRYFNMTYTYVASBGDYGTGIEAEQEARLNNICIAEAEKGRSNIRKSDS----- 258
0Y 256 RLWYTVDKLQOSTARVAVVVEFSPDLTYHFFNEVLRON---FTGAVVIASESALDPVLNL 313
Db 259 ---VIRELLQRPNAVVVLFMRSDSRRELIAAASVNASFT---WVASDGMGAQDSYK 312
0Y 314 TELGHLGFFLGITIQSVPI-----PGFSEFREMGPQAGPPPLSRTSQ 355
Db 313 SEHVAYGA-ITLELAHSHPVQEDRYFQSLNPYNHNRNPFDEWE-----Q 357
0Y 356 SYTCN-----QECNCLNATLSFNTILRSGRRVYYSVAVYVAHALHSLG--C 405
Db 358 RFQCSLQNRNRHROICDHLAIDS---NYEESKIMFVNAVAYAMAHALHKMQRTLC 412
0Y 406 DKST--CTKRVRVYPMQL--EEIMKVNFTL-----LDHQIFPDQGDVALHLEYQW 454
Db 413 PNTYLCQAMKILQCKKILYKDYLLKINFTAPNPNPKGADSYKFPYDYGMRIVWENFQ 472
0Y 455 WDRSQNPQSVASYYPLOROLKNIODISMHTVNNITPMSKSKRQSGQKKRPVGIHC 514
Db 473 HIGKYSYLYKGHW--AETLYLDVDSIHS--RNSVPTSGSDPCAPAREMKNMQPGDVC 528
0Y 515 FECIDLCFGLTNHTEDEYECQACPNNEMSYQSETSCFKROLVLEWHAEPTAVALLAA 574
Db 529 WICIPCEYEYL---VDEFTQCMDCGPGQMPADLSGCYNLBEDYIRWEDAWAIGVTTAC 585
0Y 575 LGFISTALIVFMRHPTPIYRSAGSGMCLMLTLVLVAVMNVVYVYGGPPKVSICLRO 634
Db 586 LGEFQTCIYIVFELKHNTPLPKASGRRELYTLFVGLSYCMFTFFLAKSPYICALRR 645
0Y 635 ALFPCEFTICISIAVRSFOIYCAF---KMAISREPRASVWVRQGP---YVSMATFV 687
Db 646 LGLGTSFALICYSALLTKTNCIARIEFDYKNGAQRKFTS-----PSQVFTICGLILY 698
0Y 688 LKMTIVYIGMLARPOSHRTPDDPKITIVSCNPNRNSLLENTSLDILLSVGFSEFAYM 747

```

QY	DB	Query Match	Best Local Similarity	Matches	Conservative	Mismatches	Indels	Gaps	33;
QY	14	LIM-VLAERAEKSDPYLRGDYLLGLSLH-----ANNKGIVHLNFIQVPMCKEYEVKVI 67	15.7%; Score 699.5; DB 2; Length 872;						
DB	12	LIMGVAAGRPAAKVVLTLEGGDLYLGLFVPHOKGPAEECGEPV-----EHR----- 57	26.8%; Pred. No. 2.3e-45;						
QY	68	GNLQAQARFAVEEINNDSSLPGVLLGYEIVDVCYISNN-VQPLYEFLAE-----DNL 121							
DB	58	GIQREALMFLADRIINRPHLLPGVRLGAHLILDSCKPTALHEGALDVPVRSLSRGADGS 117							
QY	122	LPIQED----YSNYISRVAVIGPDNSESVMYANFLSLFLPLQITYSAISDELDRVR 176							
DB	118	RHICPDGSVATKSDAPYATVGITGYSYDSVSIQVANILRLFIQIPQISVASTSAKLSDKSR 177							
QY	177	PALLRTTPSADHNHEANVQMLHFRMMWIIIVYSDDITGGDNQGLGERARARDICIAF 236							
DB	178	YDFEARTVPDPDFQAKAEELIRFENMTVYSTVASSEGGYGETGIEAFLEEARARNICAT 237							
QY	237	QETLPTLPQNMHTSEERQRLVTVYDKLQOQSTARVVVE--SPDLTYHFENEVLROQFT 294							
DB	238	SEKV-----GRMSAAAEQGVYRAL--LQKSAKRAVILFTSESDARELLAATQRIANST 290							
QY	295	GAVWTASESW-AIDPVLLNLTEGLHGTFLGITIOQVDPDPQSEERE---WGFOAGIPP 349							
DB	291	---VWASDQMGALSESIVAG-SERAEGA-TITELASYPIDISFASVQSLDPMNNSRNPF 345							
QY	350	LSRTQSTTCNQECNCLNATLISFTIILRSGERYVYSVYANVAVAHALSL--LGDCK 407							
DB	346	REFWEERHCGRORDCAHSL--RAVFPDESKIMFVNAVYAMAALHNRALCPN 402							

QY	408	ST C-TRRVYPMOLLEE-TWKNFLL-----LHQIFEDPOGVALHELYOW-QMD	456
Dd	403	TTTHLCDAMRPVNGRRLLKDEYLVANKFPAFPRPADTDEVAFDRDGSIGRNIFFTYIRAG	462
QY	457	RSQNPFSVASYPYLOHOLKNODISMHTVN-NTIPMSKSKRQSGOKRPVGHVCCF	515
Dd	463	SGRRYRYQVKG--YWAEGTLTDTSCFIIPASSACGLPAPRSCPELOQNEVRSVQGEJECW	520
QY	516	ECIDLOPOTFINHDEDEYEQACQNNNEMSVQSEFSCCKROLYVLEMHETIAVALLAL	575
Dd	521	LCIPIOP---KEYRLDEFTCDADGGLGYWPNASLTGCEHLEOXYIRMGDAVAVPVTTACL	577
QY	576	GFSLTALVYFNHRFOPTIYRSAGGEMPCFMTLLLLVAYVVVYVGPVPRKVSCTLCROA	635
Dd	578	GALATLFVLGVEVRHNATPVYKASGRELYTLLGGVFLGYCMFVEFLAKSPAVCTLRL	637
QY	636	LFPCLFTICICIAVRSQIYCAFKMASRPRATSVYRWQGP-----YVMAFITYLAKV	691
Dd	638	GLGTAFSYCCYALLTKTNRIARIRIGA---RECAQRPRTISPAQVAILATISGLLI	693
QY	692	IIVTIGMLARPOSHRTPDDPKITIVSCNPYRN-SLTFMTSDLLLSVGFSPAVMKE	750
Dd	694	VAAHLVVEAPBTGETAPEREVEVTLKG--NHRPASMGLATVAVLLALCTLEAFETR	751
QY	751	LPTVYNEAKFTTSLM-----TFYFTSS-----VSLCTFMSAYSGVLVTTIVDL	792
Dd	752	CPENFNEAKFIFGFMVYTTCLIMLAFLEIFYVTSDDYRVQTTMCSVSLSGSV-----	804
QY	793	LVTVLNLLAISLGVEPKCMVILPEPRNPAY	825
Dd	805	---VAGCL-----FAPRLHTILLPOPKANVSH	828

RESULT 11
A:Accession: A42916
metabotropic glutamate receptor mgluR5 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 05-Nov-1999
C:Accession: A42916
R:Abé, T.; Sugihara, H.; Nawa, H.; Shigemoto, R.; Mizuno, N.; Nakanishi, S.
J. Biol. Chem. 267, 13361-13368, 1992
A:Title: Molecular characterization of a novel metabotropic glutamate receptor mgluR5
A:Reference number: A42916; PMID:92317054; PMID:1320017
A:Accession: A42916
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1171 <AB>
A:Cross-references: GB:U010891; NID:g220813; PIDN:BAA01711.1; PID:dl0002186; PID:g220814
A:Experimental source: brain
A:Note: Sequence extracted from NCBI backbone (NCBIN:107749, NCBIPI:107750)
A:Keywords: G protein-coupled receptor; transmembrane protein

Query Match	15.4%	Score 685;	DB 2;	Length 1171;
Best Local Similarly	25.8%	Pred. NO. 4.4e-44;		
Matches 229; Conservative	152;	Mismatches 341;	Indels 166;	Gaps 32

```

QY      28 YLPDDYLLGGTSLSHANNKGIVHNLNFQVPMCKEYEVK-----ICYNLMQMRFAVEE 81
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      30 HMPDGIIGALFSVH-----H-----QPIYDKYHEKSCANVEQISIQVEMALHTLER 78
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      82 INDSLSLPGVLLGYEIVDYCYIS-----NNVQFVLYFLHEDNLILPJOEDSYTIS 133
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      79 INSDPTLLPNTLTACCEIRDSQWHSNAVALBQESIERIIRSLISSEEBGLVACVDGSSSFRS 138
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      134 R--VYANYIGPNSSSVMTVANFLSLFLPQITYSISDELKDKVRPALLRTTPSPADHHV 191
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      139 KKPITGVIGPSSSSVAIQOVNLTQLFENIPQIATSATSMDSLDTLKFYEFMRVYVPSAQQA 198
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY      192 EAMVQLMHEFMMNLIIVYSSDTGRNGQGLGRVARRDICIATFOETLPTLOPNQMS 251
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      199 RAMVDYIKRYMMYIVSAVHEGNTGSGSEMEFKMSKEGICIASHYKI-----YSN 250
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

```

0Y 252 EEPRLVITWDKOO--STRVWVWSPDITLHFENEYLRONFGA-TWISEMADLP 308
Db 251 AGEBSFKLKLKRLSHLRKRVAVACSEGTWVAGLLMARRLGLAGEFLLSDGMA--- 307
0Y 309 VLNHLTELGHGTFPLGTT--QSVPI-----DQSEFEWGHQAGP 348
Db 308 DRYVTDGYOREAVNGGITTILKQSPDVKMFEDDYILKRPETLNRPWFQEFOWHRCBLE 367
0Y 349 PLSTSGSYTNCQEDCNLNTLSFTMTILRLSERVYS---YSAYAVAHNHSILG 404
Db 368 GFAGENSKY--NRTG-----NSSJTLRTHHQDSKMGFVINALYSMAVGLHM-- 413
0Y 405 CDKSTC-----TKRVVPMOLLEEIMKVNFT--LLDQIEFDPQGDVALHLEIVQO-- 454
Db 414 -QMSLCGAYAGLDAMKPRIDGRKLDSIMKNTGVSGMILFDEMGOSPGRYTEIMNKE 472
0Y 455 -----WDSQNPQSVASTYPLQOROKIOTISHTYVNNITPMSMCKRQOS 501
Db 473 MGRDIFYINWSDMN-----ELKMDDEWVSKRNNTI-RSVCEPCEK 516
0Y 502 GQ-----RRKPVGIVHCCFECIDCLPGFLNHPEDEQOACPNNEMSYQSEISCKRQLY 557
Db 517 GQIVYIRKEYS---CCMTCTPKREMEY---FDEYTCACQOLGSPHDDLTGCOLIFVQ 570
0Y 558 FLEWHEAPITAVALLAELGLSTALIVFWRHFOPIYRSAGBPCEMLTLILVAVNY 617
Db 571 YLRMGDEPIAAVVEACLGELATLFTVYFIYIYDRPVVKSRELCTIILAGICIGYLC 630
0Y 618 VPVYVGGPRKSTCLCRQALPPLCFTTICISCIANRSQIYCAF-----KMASFPFAYSTW 672
Db 631 TFCILAPRQIYCLQNLIGISLPMSYSALVYKTRIRARILLAGSKKCTCTKPFPMSC 690
0Y 673 VRQGGPVYSAFTIV-LKAVIIVIGMLARQSHPTDPDPDKITIVSCPNPRN-SLEFN 730
Db 691 AQ-----LVIAFILICIGLQIYALIMEP-----PD-----IMHDPISIREVYLIN 733
0Y 731 TSLDLLLSVGF-----FAYMGKELPTVYNNAKFTILSMFFYFTSSVSLCTEWSA 781
Db 734 TTNGVAVTPPLGYNGLILTSCTFYAFKTRNPAFNAFNKAKIAFTM---YTCIIMLAFVPI 790
0Y 782 YSGVLAVITVLLVAVNLNLISG-YEGPCVAILVPERPRNPAYFNS 828
Db 791 YEGSNYKIITMCFVSLSATVAGCKEVPVYIILAKPERNRBSATTT 838

```

RESULT 12

JC2132
metabotropic glutamate receptor 5 A - human
C:Species: Homo sapiens (man)
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 05-Jan-1996
C:Accession: JC2132
R:Mlnakeml, R.; Katsuki, F.; Yamamoto, T.; Nakamura, K.; Sugiyama, H.
Biochem. Biophys. Res. Commun. 199, 1136-1143, 1994
A>Title: Molecular cloning and the functional expression of two isoforms of
A:Reference number: JC2131; MVID:94197696; PMID:7908515
A:Accession: JC2132
A:Molecule type: mRNA
A:Residues: 1-1180 <MIN>
A:Comment: mkbv92091c cloned from human neuroblastoma, binds serotonin

C:Comment: This protein is coupled to guanine nucleotide binding proteins
C:Keywords: glycoprotein; neurotransmitter; receptor; transmembrane protein
F:580-604/Domain: transmembrane #status predicted <TM1>
F:617-637/Domain: transmembrane #status predicted <TM2>
F:644-664/Domain: transmembrane #status predicted <TM3>
F:654-714/Domain: transmembrane #status predicted <TM4>
F:738-759/Domain: transmembrane #status predicted <TM5>
F:773-794/Domain: transmembrane #status predicted <TM6>
F:803-827/Domain: transmembrane #status predicted <TM7>

Query Match	15.2%;	Score 675.5;	DB 2;	Length 1180;
Best Local Similarity	25.7%;	Pred. No. 2.4e-43;		
Matches 230;	Conservative 153;	Mismatches 333;	Indels 179;	Gaps 34;
Oy	28	YLPDGYLLGGLESLHANKKGIYHNFLOYPCKEYEVK-----IGYVLQAMRPAVDE	81	


```

Db      30 HMPGDIIGALFVSF-----H-----OPTVDKVERKCGAVREOYGIQREAMLHTLER 78
      82 INNDSSLPGVLLGEIYDVCYSN-NVQPVLYFLAHEDNLLPIQED-----YSN 130
      79 INSDPTLLPNTLLGCEIRDSCWHSVALEOSTIEFT-RDLSISSEEGGLVRCVDGSSSS 136
      131 YISR--VVAVIGPDNSESMTVANFLSLFLPOLITYSAISDELBRKVPFALLRTTPSAD 188
      137 FRSKRPVIGVIGPSSVAIQVNLQLEPNIPQIAVSATSMDSKLTLEKFMRYVPSDA 196
      189 HVEAMVOLMLHFRRNMWIIIVSSDPTGRCNGQLGERVARDICIAFOELPTLOPNON 248
      197 QOARMAVDIVKRYNMTYSAVHTBEGNYESGMEAPKDSAKGICIAHSYKRYSNAGSOS 256
      249 MTSEERORLVTIVDKLOOS-----TARVVYFSPDLTLHYEFNEVLKRONFTGA-VWIAS 301
      257 F-----DKLLKLTSHLPKARVAVACFCGKMTVRGLLMMRRLLGLAGEFLILGS 304
      302 ESWAIDPVLHNTLGHGTLGTLTI--QSVPI-----PGSESEFREW 341
      305 DQMA---DRYDVTGQYREAVAGGITIKLOSPOVKWFDYTLKRPETHNRNPFQEFWOH 361
      342 GPQAGPPLSRTSOSYTCNOECDNCLNATLSNTILRLSGERYYS---YSAVAYAAH 397
      362 RFQCRLEGPPOENSKY--NKTC-----NSSLTLEKTHVODSKMGFYINATYSMAV 409
      398 ALHSLGCDKSTC-----TKRVVYPMOLLEIEIMKVNFT-LDHOQTFPDQGVALLH 448
      410 GLHNN---QMSLCPGYAGLCDAKMRPIDGRKLESLMKTNFVGSODTLIFDENGSPGRY 466
      449 EIVOMQ-----MDRSQNPQSVASVYPLOROLKNIDISWHTVNNTPISM 494
      467 EIMNFKEMGKDYFDYINGSMONG-----ELKMDDEWMSKSNII--RSV 510
      495 CSKRQSGO---KKRPVGIHVCCPECIDCLPGFTLNTHEDEYECQACPNNMWSQSETS 550
      511 CSEPEKQIKVIRKGEVS---CQWTCPPCKENEYV---FDEYTCACQOLGSMPTDILTG 564
      551 CEKROLVLEWHEAPTIIVALLAAGFLSTALIVIFMRHPTPIVRSAGGPMCLMTL 610
      565 CDLIPOVQILRMGDPEPIAAVAVPACIGLATLFTVYVFIIRPTPVKSSRELCTIILAG 624
      611 LLVAYMVVPPVYVGPVKVSTCLCROALFPLCTICISCIASVRSQIVCAF---FMASRF 665
      625 ICLGILCTFCLAKKQIYCYLQIRIGISLSPMYSALVTKNRIARILAGSKKICRK 684
      666 PRAYSVWRYOGPYVSMATFY-LKMWIVVIGMLARPOSHPTDDDKITIVSCNPNYR 724
      685 PRFMSACQO---LVIAFILICIQIGITVAFIMEP-----PD-----IMHDYPSIR 727
      725 N-SLFLNLSLDLLSVGES-----FAYMGKELPTNYNEAFITLSMFTYTSVS 774
      728 EYLLICNTNNGVVPPLGNGILLISCIFYAKRTNVANFNAFTYIAFTM---YTTGCI 784
      775 ICTFMSATSGVLTIVDLLVTVNLTAISLG-YFGPKCYMLFYPERNTPAYFNS 828
      785 WLAFVPIYFGSNYKIITIMCFSVLSATVALGCMFVPKYIILIAKERVRSAFIT 839

RESULT 13
JC2131
metabotropic glutamate receptor 5 B - human
C:Species: Homo sapiens (man)
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 05-Jan-1996
C:Accession: JC2131
R:Minakami, R.; Katsuki, F.; Yamamoto, T.; Nakamura, K.; Sugiyama, H.
Biochem. Biophys. Res. Commun. 199, 1136-1143, 1994
A:Title: Molecular cloning and the functional expression of two isoforms of human metabo
A:Reference number: JC2131; MUID:94197696; PMID:7908515
A:Accession: JC2131
A:Molecule type: mRNA
A:Residues: 1-1212 <MIN>

```

```

C:Comment: This protein is coupled to guanine nucleotide binding proteins.
C:Keywords: glycoprotein; neurotransmitter; receptor; transmembrane protein
F:580-604/Domain: transmembrane #status predicted <TM1>
F:617-637/Domain: transmembrane #status predicted <TM2>
F:644-664/Domain: transmembrane #status predicted <TM3>
F:694-714/Domain: transmembrane #status predicted <TM4>
F:738-759/Domain: transmembrane #status predicted <TM5>
F:773-784/Domain: transmembrane #status predicted <TM6>
F:803-827/Domain: transmembrane #status predicted <TM7>

Query Match      15.2%; Score 675.5; DB 2; Length 1212;
Best Local Similarity 25.7%; Pred. No. 2,5e+43;
Matches 230; Conservative 153; Mismatches 333; Indels 179; Gaps 34;

      28 YLPGDYLLGLGFLSHANKGIVHNLFOVPMCKEYEVY-----IGYNMQAMRFAVEE 81
      30 HMPGDIIGALFVSF-----H-----OPTVDKVERKCGAVREOYGIQREAMLHTLER 78
      82 INNDSSLPGVLLGEIYDVCYSN-NVQPVLYFLAHEDNLLPIQED-----YSN 130
      79 INSDPTLLPNTLLGCEIRDSCWHSVALEOSTIEFT-RDLSISSEEGGLVRCVDGSSSS 136
      131 YISR--VVAVIGPDNSESMTVANFLSLFLPOLITYSAISDELBRKVPFALLRTTPSAD 188
      137 FRSKRPVIGVIGPSSVAIQVNLQLEPNIPQIAVSATSMDSKLTLEKFMRYVPSDA 196
      189 HVEAMVOLMLHFRRNMWIIIVSSDPTGRCNGQLGERVARDICIAFOELPTLOPNON 248
      197 QOARMAVDIVKRYNMTYSAVHTBEGNYESGMEAPKDSAKGICIAHSYKRYSNAGSOS 256
      249 MTSEERORLVTIVDKLOOS-----TARVVYFSPDLTLHYEFNEVLKRONFTGA-VWIAS 301
      257 F-----DKLLKLTSHLPKARVAVACFCGKMTVRGLLMMRRLLGLAGEFLILGS 304
      302 ESWAIDPVLHNTLGHGTLGTLTI--QSVPI-----PGSESEFREW 341
      305 DQMA---DRYDVTGQYREAVAGGITIKLOSPOVKWFDYTLKRPETHNRNPFQEFWOH 361
      342 GPQAGPPLSRTSOSYTCNOECDNCLNATLSNTILRLSGERYYS---YSAVAYAAH 397
      362 RFQCRLEGPPOENSKY--NKTC-----NSSLTLEKTHVODSKMGFYINATYSMAV 409
      449 EIVOMQ-----MDRSQNPQSVASVYPLOROLKNIDISWHTVNNTPISM 494
      467 EIMNFKEMGKDYFDYINGSMONG-----ELKMDDEWMSKSNII--RSV 510
      495 CSKRQSGO---KKRPVGIHVCCPECIDCLPGFTLNTHEDEYECQACPNNMWSQSETS 550
      511 CSEPEKQIKVIRKGEVS---CQWTCPPCKENEYV---FDEYTCACQOLGSMPTDILTG 564
      551 CEKROLVLEWHEAPTIIVALLAAGFLSTALIVIFMRHPTPIVRSAGGPMCLMTL 610
      565 CDLIPOVQILRMGDPEPIAAVAVPACIGLATLFTVYVFIIRPTPVKSSRELCTIILAG 624
      611 LLVAYMVVPPVYVGPVKVSTCLCROALFPLCTICISCIASVRSQIVCAF---FMASRF 665
      625 ICLGILCTFCLAKKQIYCYLQIRIGISLSPMYSALVTKNRIARILAGSKKICRK 684
      666 PRAYSVWRYOGPYVSMATFY-LKMWIVVIGMLARPOSHPTDDDKITIVSCNPNYR 724
      685 PRFMSACQO---LVIAFILICIQIGITVAFIMEP-----PD-----IMHDYPSIR 727
      725 N-SLFLNLSLDLLSVGES-----FAYMGKELPTNYNEAFITLSMFTYTSVS 774
      728 EYLLICNTNNGVVPPLGNGILLISCIFYAKRTNVANFNAFTYIAFTM---YTTGCI 784
      775 ICTFMSATSGVLTIVDLLVTVNLTAISLG-YFGPKCYMLFYPERNTPAYFNS 828
      785 WLAFVPIYFGSNYKIITIMCFSVLSATVALGCMFVPKYIILIAKERVRSAFIT 839

```


RESULT 14
A46742
metabotropic glutamate receptor, mGluR6 - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 20-Jun-2000
C:Accession: A46742
J: Nakajima, Y.; Iwakabe, H.; Akazawa, C.; Nawa, H.; Shigemoto, R.; Mizuno, N.; Nakanishi, J. *Biol. Chem.* 268, 11868-11873, 1993
A:Title: Molecular characterization of a novel retinal metabotropic glutamate receptor
A:Reference number: A46742; MUID:93280152; PMID:8389366
A:Accession: A46742
A>Status: preliminary
A:Molecule type: nucleic acid
A:Residues: 1-871 <NA>
A:Cross-references: GB:D13963; NID:g391856; PIDN:BA03066.1; PID:g391857
A:Experimental source: retina
A>Note: sequence extracted from NCBI backbone (NCBI:133246, NCBI:P133250)
C:Superfamily: metabotropic glutamate receptor 4
C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 15.1%; Score 673; DB 2; Length 871;
Best Local Similarity 26.4%; Pred. No. 2,5e-43;
Matches 239; Conservative 155; Mismatches 377; Indels 134; Gaps 33;

OY 11 LEFLWLAE--DAENSDFYLPDGYLLGLFSLHANNKGIHNLNFOVPMCKEYEVKVI 67
DB 8 LMLMLWLSQAGIACGAGSVLAGLGLFVPHARGAAGACGALK-----KEQ 58
OY 68 GYNLMQMRFAVEIINDSLLPVLGLEYVDVCIYISNNVQPVLYFLAHEDNLLPIQED 127
DB 59 GVHRLKMLALDHYVNDPELLPVGRLGARLLDTC--SRDYALQALSFQVALIRGRGD 116
OY 128 -----YSNYISRVAVAVIGPDNSESVMTVANLSLFLPOLITYASIDELR 172
DB 117 GDEASVRCGCVPLRSPRRVAVVAGASSVISVAVANLRLFAIPQISTASTABELS 176
OY 173 DKVFPALLRTTPSADHHEAMVOLMLHFRWNWIIIVSSDPTGRDNGQLGEEVARR-- 230
DB 177 DSTRIDEFSSRVVPPDSYQAOAMDIVRALGMVYSTLASSENG-ESGVEAFVQISREAG 235
OY 231 DICIAFOETLPTLPONQMTSEEROLVYDKLOQSTARVVVPSDLTIHFENEYLR 290
DB 236 GVCIAQOSTIKIPR-EPRGEEFKVIRL-----METPNARQIIIFANEDDIRVLEATRO 288
OY 291 QNFTG-AVMIASESMA--IDPVLNLTLE--LGHGTFLGII--QSVPIPFSEFRWGP 343
DB 289 ANLGHFLWAGSDSGWSKISPL-LLEEAVG-----AIIILPKRASIDGFDQYF---- 337
OY 344 QAGPPLSRTSOSTYCN-----QECDCNCL--LNTLSFNTILRLSGER----- 383
DB 338 -----MTRSLNNRRNRINWFAEFEEENFNCKLTSSSGOSDSTFKCKEERIGODSAVEQ 391
OY 384 --VYYSVYSAVVAHAHLHL--LGCDKSCCTKRIVVYPMW--LLEELWVYNFT-LLDHQ 435
DB 392 EGKVOYIADAYATAHLSHQAQLCPGHGTGLCPAMEPTDGRITLHYTRAENGSGAGTP 451
OY 436 IFDPQDVALHLEIYQWQ--MDRSONPQSVASYPLQROLKNIDISWHYNNNTIPM 492
DB 452 VMFENDADAGRVDIPOYQATNSGASSGGYQAVGQAMALR--LDMVLKRSQDPHEVPP 509
OY 493 SMCSKRCOSQOKKPVGIHVCCFECIDCLPGLNHTEDDEYEQACPNNMYSQSETSCF 552
DB 510 SOCLSPGCPGRKKRVKGVPCMCCEACDGYRF--QVDEPTCEACGDMRPTNHTGCR 566
OY 553 KRQVLEFHEHAPPIAVALAALGFLSTALIVTFMRHFOPTPIYRSAGGPMCFMLTL 612
DB 567 PTPVVRLLTWSSPMALPDLALLAVIGIMATYIMAFMRHNDPTPIYRASGRELSYLLTGIF 626
OY 613 VAVVAVVYVVPKSTCROALFPLCFITICICIAVRSPQIYCAKMAKSRFPRAVSVM 672
DB 627 LITAITFLMAEFCALICARRLLIGITLISALTITKIRIYRIFEGCKRSVT----- 681

OY 673 VRYQPYVS-----MAF-ITVLMKVIYVIGMLARP-----QSHRTPDDPKITIVS 718
DB 682 ---PPFISPTSQLVITFGGLSLQVAVIAMLGAQPPHSVIDVEQRTVDPEQAR-GVLR 737
OY 719 CNPNYRNSLLPNTSLDLLSVGSEFAYMGKELPTYNNAKFTITLSRTYFTSSVSLCTF 778
DB 738 CDMS-DLSLGLCYGLSLLLVCTVVAIKARGVPEFENKAPIGFTW---YTCIIWLAF 793
OY 779 MSFVSG-----VLVYIVDLVTVLNLALISIG-YRPRCYWILFPERNTPAFENSMI 830
DB 794 VPVFPGIAQNAEKIYIQTITLVISLSASVSLGMLVPTTYITLHPEDQVOKRRSLK 853
OY 831 QGYTM 835
DB 854 KTSFM 858

RESULT 15
I49142
metabotropic glutamate receptor 8 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C:Accession: I49142
R: Duvoisin, R.M.; Zhang, C.; Ramonell, K.
J. *Neurosci.* 15, 3075-3083, 1995
A:Title: A novel metabotropic glutamate receptor expressed in the retina and olfactory
A:Reference number: I49142; MUID:95239344; PMID:7722646
A:Accession: I49142
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-908 <RES>
A:Cross-references: EMBL:U17252; NID:g854728; PIDN:AAA68149.1; PID:g854729
C:Genetics:
A:Gene: mGluR8
C:Superfamily: metabotropic glutamate receptor 4
C:Keywords: neurotransmitter receptor

Query Match 15.1%; Score 672; DB 2; Length 908;
Best Local Similarity 25.2%; Pred. No. 3.1e-43;
Matches 220; Conservative 176; Mismatches 375; Indels 132; Gaps 37;

OY 4 RAKTICSLFLL-----WLA--EPAENSDF-----YLPDYLGLGFLSLHA-NMKGIYHL 51
DB 6 KRSTSCCPFLAKFAFWILTMQRTSHSOYASIRLDGIIIGGLFPHAKERGV-- 62
OY 52 NFLQVPMCKEYEVKVIYGNLMQMRFAVEIINDSLLPVLGLEYVDVCIYISNNVQPV 111
DB 63 -----PCGDLK-KEGIIHLEMLYAIIDQTNMDPOLLSTITIGVRLDTC--SRDYAL 113
OY 112 LVFLAHEDNLLPIQEDYSNT-----ISRVAVAVIGPDNSESVMTVANFLSLFL 159
DB 114 EQSLTFVQAL--LEKQASDVKANGDPPIFTKPKDISVIGGAASSVIMAVANLRLFL 171
OY 160 POLITYASIDELNDKVFAPLRTTPSADHHEAMVOLMLHFRWNWIIIVSSDPTGRDN 219
DB 172 POLISTASTABELDNTRYDEFSSRVVPPDSYQAOAMDIVRALGMVYSTLASSENG-ES 230
OY 220 GOLLEGEVARR--DICIAFOETLPTLPONQMTSEEROLVYDKLOQSTARVVVFPSP 277
DB 231 GVAFQIISREIGVCIQAQOKIPR-EPRGEEFKIKRL-----LETNAAAVIMEAN 283
OY 278 DLTLYHFENFVLRONTG-AVMIASESMA--IDPVLNLTLELGHGTFLGII--QSVPI 332
DB 284 EDIIRGLLEAKKLNGSHGLWIGSDSGWSKIAPV-YQGEIAE-----GAVTILPKRASI 338
OY 333 PG-----FSEPRE--WGPQAGPPLSTSYCQNDCCNLANNTLS 372
DB 339 DGEFDRYRSRTLANNRNWFAEFSENGFCCKSGSHG-KRNSIRKCT-----G 386
OY 373 FNTILRLSG---ERVYYSVYSAVVAHAHLHLGLGDKSCTKRV-----VYPMOLL 421
DB 387 LERIAQSYEQBQKVOFVIDAVYSMAVYALHNM--HKEICPGYIGLCPRMTTIDGKELL 443

OY 422 BEIWKVNT-LIDHOIFEDPOGDVALHLEIYOMQMDRSONPROQSVASYYPLOROLKNIOD 480
| | | | :
Db 444 GYIRAVNNGSAGTPEVTNENGADGRYDIFQYQINNKSTYK-TIGHWNOHLK-VED 501
OY 481 ISWHTVNTTIPMSCKRCQSGOKKRPVGIHVCCFECIDCLPGTEPLNTEDEYECQACPN 540
:
Db 502 MQMANREHTHPASVCSLPCFKPGERKKTVKGVCCWHGRC---EGYNYQVDELSCELCPL 558
OY 541 NEMSQSTSCFKROLVLEHHEAFTIYVALLAIGFLSTLAILVIFWRHROTPIYRSAG 600
:
Db 559 DQRPINRTGQRPILKLEHSPWAVVPVLAIGIATTEVIYTFVRYNDTPIYRASG 618
OY 601 GPMCFIMLTLLVAVMVPVYVGPVKVSTCICROALPFLCFTICISCIJAVRSFOIYCAFK 660
:
Db 619 RELSTVLTGIFLCYSTIFLMIAPDTIICSPRRIFLGMCFSYAALLITKNIHRIPE 678
OY 661 MASRPPRAYSYWRYQGPVSNMFTVLKMYIVVIGMLARQS---HPRT-DEDDPKI 714
:
Db 679 QGKKSVPAPKFISPASQLVITFSLISYQLGVFWFVVDPPHTIIDYGEORTLDPENAR- 737
OY 715 TIVSCNPNYRNSLENTSLDLLSVGSEFAYMGKELPTYNNEAKFTLSMTFTSSYS 774
:
Db 738 GVLKCDIS-DLSLIGSLGYSILMWTCTVYAIKTRGVDETFENAKPIGFTM---YTCII 793
OY 775 LCTFMSAYSQ-----VLVTIVDLTVLNLNLAISIG-YFGPKCYMILEYPERNT---P 823
:
Db 794 WIAFTIPFPGTAQSAEKMYIQTTTLTVSMSLSASVSLGMLYMKYIIIIHFPEQNVQKRK 853
OY 824 AYFNSMIOGYTMR 836
:
Db 854 RSFKAVYTAATMQ 866

Search completed: May 19, 2003, 09:51:13
Job time : 24.8343 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.P5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:46:50 ; Search time 10.9089 Seconds
(without alignments)
3186.136 Million cell updates/sec

Title: US-09-927-315-9
Perfect score: 4443
Sequence: 1 MGRPACTICLFLFWLAE.....ERNTPATFNSMIGYIMRD 838

Scoring table: BIOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues
Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1094	24.6	1078	1 CASR_HUMAN	P41180 homo sapien
2	1088.5	24.5	1079	1 CASR_RAT	P48442 rattus norv
3	1087.5	24.5	1085	1 CASR_BOVIN	P35384 bos taurus
4	1085.5	24.4	1079	1 CASR_MOUSE	O99Y96 mus musculu
5	724.5	16.3	915	1 MGR7_RAT	P35400 rattus norv
6	721.5	16.2	912	1 MGR4_RAT	P31423 rattus norv
7	721.5	16.2	915	1 MGR7_HUMAN	Q14831 homo sapien
8	718	16.2	879	1 MGR3_RAT	P31422 rattus norv
9	717.5	16.1	912	1 MGR4_HUMAN	Q14833 homo sapien
10	709	16.0	877	1 MGR3_HUMAN	Q14832 homo sapien
11	699.5	15.7	872	1 MGR2_RAT	P31421 rattus norv
12	689	15.5	976	1 MGR_DROME	P31685 drosophila
13	686.5	15.5	872	1 MGR2_HUMAN	Q14416 homo sapien
14	685	15.4	1203	1 MGR5_RAT	P31424 rattus norv
15	675.5	15.2	1212	1 MGR5_HUMAN	P41594 homo sapien
16	674.5	15.2	877	1 MGR6_HUMAN	O15303 homo sapien
17	673	15.1	871	1 MGR6_RAT	P35349 rattus norv
18	673	15.1	908	1 MGR8_RAT	P70579 rattus norv
19	672	15.1	908	1 MGR8_MOUSE	P47743 mus musculu
20	666.5	15.0	908	1 MGR8_HUMAN	O00222 homo sapien
21	655.5	14.8	999	1 MGR1_CAEL	O09630 caenorhabdi
22	654	14.7	1199	1 MGR1_RAT	P33385 rattus norv
23	647	14.6	1194	1 MGR1_HUMAN	Q13255 homo sapien
24	180.5	4.1	960	1 GBR1_MOUSE	Q94W18 mus musculu
25	177.5	4.0	961	1 GBR1_HUMAN	Q94B45 homo sapien
26	169	3.8	991	1 GBR1_RAT	O92044 rattus norv
27	147.5	3.3	940	1 GBR2_RAT	O88871 rattus norv
28	144	3.2	918	1 GIK1_HUMAN	P39086 homo sapien
29	144	3.2	949	1 GIK1_RAT	P22756 rattus norv
30	140.5	3.2	836	1 GIK1_MOUSE	O60934 mus musculu
31	139.5	3.1	941	1 GBR2_HUMAN	O75899 homo sapien
32	131.5	3.0	1125	1 CYGS_STRPU	P16065 strongyloce
33	130.5	2.9	938	1 NM2L_HUMAN	Q05586 homo sapien

ALIGNMENTS

RESULT 1	CASR_HUMAN	STANDARD	PRT: 1078 AA	
AC	P41180: Q13912; Q16379; Q16108; Q16109; Q16110;			P35439 rattus norv
DT	01-FEB-1995 (Rel. 31, Created)			P35438 mus musculu
DT	01-NOV-1995 (Rel. 32, Last sequence update)			P34352 caenorhabdi
DT	15-JUN-2002 (Rel. 41, Last annotation update)			Q94126 homo sapien
DE	Extracellular calcium-sensing receptor precursor (Casr) (Parathyroid Cell calcium-sensing receptor).			P19490 rattus norv
GN	CASR OR GPRC2A OR PCAR1.			P23818 mus musculu
OS	Homo sapiens (Human).			O24738 drosophila
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteria; Primates; Carnivora; Homiidae; Homo.			P34976 oryctolagus
OX	NCBI_Taxid:9606;			P45443 neotoma
NP	SEQUENCE FROM N.A.			P20594 homo sapien
RA	Pearce S.H.S., Thakker R.V.;			P51841 homo sapien
RL	Submitted (Dec-1994) to the EMBL/GenBank/DBJ databases.			P30555 sus scrofa
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Parathyroid;			
RX	MEDLINE=95279439; PubMed=7759551;			
RT	Garrett J.E., Capuano I.V., Hammerland L.G., Hung B.C., Brown E.M., Hebert S.C., Nemeth E.F., Fuller F.;			
RL	"Molecular cloning and functional expression of human parathyroid calcium receptor cDNAs.";			
RN	J. Biol. Chem. 270:12919-12925(1995).			
RC	[3]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Kidney;			
RX	MEDLINE=95408281; PubMed=7677761;			
RA	Aida K., Koishi S., Tawata M., Onaya T.;			
RT	"Molecular cloning of a putative Ca(2+)-sensing receptor cDNA from human kidney.";			
RL	Biochem. Biophys. Res. Commun. 214:524-529(1995).			
RN	[4]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96343808; PubMed=8756555;			
RT	Fretschel M., Zink-Lorenz A., Hollloschi A., Hafner M., Flockert V., Rate F.;			
RL	"Expression of a calcium-sensing receptor in a human medullary thyroid carcinoma cell line and its contribution to calcitonin secretion.";			
RN	Endocrinology 137:3842-3848(1996).			
RP	[5]			
RX	SEQUENCE OF 643-908 FROM N.A.			
RA	MEDLINE=96193893; PubMed=8613532;			
RT	Bikle D.D., Ratnam A., Mauro T., Harris J., Pillai S.;			
RL	"Changes in calcium responsiveness and handling during keratinocyte differentiation. Potential role of the calcium receptor.";			
RN	J. Clin. Invest. 97:1085-1093(1996).			
RP	[6]			
RX	VARIANTS FHH GLN-185; LYS-297 AND TRP-795.			
RA	MEDLINE=94094324; PubMed=7916660;			
RT	Pollak M.R., Brown E.M., Chou Y.-H.W., Hebert S.C., Marx S.J., Steinmann B., Levi T., Seidman C.E., Seidman J.G.;			

RT "Mutations in the human Ca(2+)-sensing receptor gene cause familial
RT hypocalcemic hypercalcaemia and neonatal severe
RT hyperparathyroidism." ;
RN Cell 75:1297-1303(1993).
RN [7]
RP VARIANT ADH ALA-127.
RX MEDLINE=95119179; PubMed=7874174;
RA Pollak M.R., Brown E.M., Estepe H.L., McLaine P.N., Kilfor O., Park J.,
RT Hebert S.C., Selman C.E., Selman J.G.;
RT "Autosomal dominant hypocalcaemia caused by a Ca(2+)-sensing receptor
RT gene mutation." ;
RL Nat. Genet. 8:303-307(1994).
RN [8]
RP VARIANTS FHH MET-62; CYS-66; MET-138; GLU-143 AND GLN-227.
RX MEDLINE=95243222; PubMed=7726161;
RA Chou Y.-H.W., Pollak M.R., Brandi M.L., Toss G., Argyrakis H.,
RA Atkinson A.B., Papapoulos S.E., Marx S., Brown E.M., Selman J.G.,
RA Selman C.E.;
RT "Mutations in the human Ca(2+)-sensing receptor gene that cause
RT familial hypocalcemic hypercalcaemia." ;
RL Am. J. Hum. Genet. 56:1075-1079(1995).
RN [9]
RP SEQUENCE OF 1-61 FROM N.A., AND VARIANT FHH ALA-39.
RX MEDLINE=95403641; PubMed=7673400;
RA Aida K., Koishi S., Inoue M., Nakazato M., Tawata M., Onaya T.;
RT "Familial hypocalcemic hypercalcaemia associated with mutation in the
RT human Ca(2+)-sensing receptor gene." ;
RL J. Clin. Endocrinol. Metab. 80:2594-2598(1995).
RN [10]
RP VARIANTS NSHPT LEU-227 AND TYR-598.
RX MEDLINE=96292293; PubMed=8675635;
RA Pearce S.H.S., Trump D., Wooding C., Besser G.M., Chew S.L.,
RA Grant D.B., Heath D.A., Hughes I.A., Paterson C.R., Whyte M.P.,
RA Thakker R.V.;
RT "Calcium-sensing receptor mutations in familial benign hypercalcaemia
RT and neonatal hyperparathyroidism." ;
RL J. Clin. Invest. 96:2683-2692(1995).
RN [11]
RP VARIANTS ADHP THR-116; HIS-681 AND SER-806, AND VARIANT SER-851.
RX MEDLINE=96311554; PubMed=8733165;
RA Baton J., Whiner K.K., Yanovski J.A., Cunningham A.W., Laue L.,
RA Zimmerman D., Cutler G.B. Jr.;
RT "Mutations in the Ca(2+)-sensing receptor gene cause autosomal
RT dominant and sporadic hyperparathyroidism." ;
RL Hum. Mol. Genet. 5:601-606(1996).
RN [12]
RP VARIANT FHH ARG-174.
RX MEDLINE=97442275; PubMed=9298824;
RA Ward B.K., Stuckey B.G.A., Gutteridge D.H., Laing N.G., Pullan P.T.,
RA Ratajczak T.;
RT "A novel mutation (L174R) in the Ca2+-sensing receptor gene
RT associated with familial hypocalcemic hypercalcaemia." ;
RL Hum. Mutat. 10:233-235(1997).
RN [13]
RP VARIANT FHH GLU-557.
RX MEDLINE=91603857; PubMed=11762699;
RA Nakayama T., Minato M., Nakagawa M., Soma M., Tohe H., Aoi N.,
RA Kosuge K., Sato M., Ozawa Y., Kamatsuse K., Kokubun S.;
RT "A novel mutation in Ca2+-sensing receptor gene in familial
RT hypocalcemic hypercalcaemia." ;
RL Endocrine 15:277-282(2001).
CC -1- FUNCTION: SENSE CHANGES IN THE EXTRACELLULAR CONCENTRATION OF
CC CALCIUM IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A
CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
CC MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
CC produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: FOUND IN KIDNEY, BUT NOT IN BRAIN, LUNG,
CC LIVER, HEART, SKELETAL MUSCLE, OR PLACENTA.
CC -1- DISEASE: DEFECTS IN CASR ARE A CAUSE OF FAMILIAL HYPOCALCAEMIC
CC HYPERPARAETHYROIDISM (FHH) AND NEONATAL SEVERE HYPERPARATHYROIDISM
CC (NSHPT). TWO INHERITED CONDITIONS CHARACTERIZED BY ALTERED CALCIUM

[illegible]

Query Match 24.5%; Score 1088.5; DB 1; Length 1079;
 Best Local Similarity 31.7%; Pred. No. 7.5e-73;
 Matches 284; Conservative 164; Mismatches 328; Indels 121; Gaps 31;

FT CARBOHYD 468 468 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 488 488 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 541 541 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 594 594 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1079 AA; 120867 MW; D76645503619736 CRC64;

DB 9 ALLALAW-----HSSAYGPDORAKKGDILIGLPIH--GVAKD--QDKSRPE 56
 DB 63 EVKIVGYNL-----MOARFAVEEINNDSSLLPGVLLGEYELVDYC-ITSNVQVLYEIA 116
 DB 57 SVECTIRNFRGFWLOAMIFAEIENSSPLIPMTIGRIFDCNYSKALEITLSVA 116
 DB 117 HEDNLPJIOE--DYSNYSRVAVIGPDNSVMTANFLSLPQITYSATSDERD 173
 DB 117 ONKIDSLNDFCNCSEHIPSTIAVVGATSGSTAVANILGLFYIPQVSYASSRLSN 176
 DB 174 KVRPALLRTPSADHVEANVQMLFRMWITIVLVSDTYGDNQGLGERVARDIC 233
 DB 177 KNOYKSFRTIPNDHATAMADIEFRMNVGTIAADYDGRPGIEKFEKREAEERDIC 236
 DB 234 IAFQETPLTLPONMTSEERQRLVTVYDKLOQSTARVYVSPDLTYHFENVLKRF 293
 DB 237 IDFESELT-----SOYSDDEEIQVVEYI--QNSTAKYIVFSSGPDLEPIKRIEIVRNI 288
 DB 294 TGAIVASSEMADIPVNLNTEIGH-IGTFIGITIOSVPIPFSEF----- 338
 DB 289 TGRIMLSSEAWASSLI-AMPEYFHVVGCTIGFGIKAGQIDFREFEFLQVPRKSVHNGF 347
 DB 339 -RENGPOA-----GPPPLSTSGSYTCNOECNCLNATLSEPTILSGE----- 382
 DB 348 AKETWEETFNCHLOBGAQKPLPVTFYRSH--EGCGNRLNSSTAFRPL--CTGDENINS 403
 DB 383 -----RVVYSYSAVAYAAVAHALSHLAC-----DKSTCKRVVYPMOLLEE 423
 DB 404 VEFYMYDEHLRISYNYLLAVYSIAHMLQDLYTCLPGRGLTNSCADIKKVEAMOVK 463
 DB 424 IMKKNFT-LDHOIFPPDQGVAAHLELVOMQDRSONP--FQSYASTYPL-----QROL 475
 DB 464 LRHLNFTNNMGEQVTFECCGLVGNYSIIMHLSPEDESIYFKEV-GYNNYAKKGETLF 522
 DB 476 KNIDISHTVNTIPMSGKRCOSGOKKPV-GIHVCECECIDCLPGLNTHEDYE 534
 DB 523 INEKILMSGFREVPSNCSRDQAGTRKGIIGSEPTCECEVECPDGEISGET-DASA 581
 DB 535 COACPNNMWSYQSEFCFRQVLEFHEHAPTIAVALLAALGFLSTIAIVTFMRHPTP 594
 DB 582 CDKCPDPMFMSNENTSCIAKIEFLAWTEPGIALTLFAVIGFLTAVLGVFIKFRNTP 641
 DB 595 IYRAGSGPMCLMTLLIVAVMVVYVGPVKYSTICRQALFPLCFITICSLAVNSFQ 654
 DB 642 IVKATNELSLTLLESLCCSSSLFPIGEPODMTCRKOPAFGISFLCLISCLIVKTR 701
 DB 655 IVCAF--KMASRPPRAVSYYRQGYVSMAFITVLKAVIYVIGMLARPOSHPRTPDDP 712
 DB 702 VLLFPEAKIPISFHRK--WGLMLQPLVFLCTFMOLITIIIVYAPBSYNNHEHD 758
 DB 713 KITIVSCPNYRNSLFTSL--DLLSYVGSFAVMGKELPTNYNAKETITLSMTYF 769
 DB 759 EIIPTYC--HEGLMAGSLIGYTCLLAICFPFAFRSKRLPENPNNAKTITSLIFF 815
 DB 770 TTSVSLCTFMSAGCVLTIVDLLVYVNLALISLG----YFGKCIWILFPERNT 822
 DB 816 IWMIS--FIPAYASTYGEKFS-AVEVAILAASFGILACIFENKVVYIILFKPSRNT 868

ID CASR_BOVIN STANDARD; PRT; 1085 AA.
 AC P35384;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Extracellular calcium-sensing receptor precursor (CasR) (Parathyroid
 DE cell calcium-sensing receptor).
 GN CASR OR GPRC2A OR PCAR1.
 OS Bos taurus (Bovine).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
 OC Bovidae; Bovine; Bos.
 OC NCBI_TaxID=9913;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Parathyroid;
 RX MEDLINE=94077162; PubMed=8255296;
 RA Brown E.M., Gamba G., Riccardi D., Lombardi M., Butters R., Klfor O.,
 RA Sun A., Hediger M.A., Lytton J., Hebert S.C.;
 RT "Cloning and characterization of an extracellular Ca(2+)-sensing
 RT receptor from bovine parathyroid."
 RL Nature 366:575-580(1993)
 CC -!- FUNCTION: SENSE CHANGES IN THE EXTRACELLULAR CONCENTRATION OF
 CC CALCIUM IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; S67307; AAB29171.1; -;
 DR PIR; S40476; S40476.
 DR InterPro; IPR001828; ANF_receptor.
 DR InterPro; IPR000337; GPCR_Mgr.
 DR Pfam; PR00003; 7tm_3; 1.
 DR Pfam; PF01094; ANF_receptor; 1.
 DR PRINTS; PR00248; GPCRMR.
 DR PROSITE; PS00979; G_PROTEIN_RECPEP_F3_1; 1.
 DR PROSITE; PS00980; G_PROTEIN_RECPEP_F3_2; 1.
 DR PROSITE; PS00981; G_PROTEIN_RECPEP_F3_3; 1.
 DR PROSITE; PS00259; G_PROTEIN_RECPEP_F3_4; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 1085
 FT DOMAIN 20 613
 FT TRANSSEM 614 636
 FT DOMAIN 637 650
 FT TRANSSEM 651 671
 FT DOMAIN 672 682
 FT TRANSSEM 683 701
 FT TRANSSEM 702 725
 FT TRANSSEM 726 746
 FT DOMAIN 747 770
 FT TRANSSEM 771 793
 FT TRANSSEM 794 806
 FT TRANSSEM 807 829
 FT DOMAIN 830 837
 FT TRANSSEM 838 863
 FT DOMAIN 864 1085
 FT CARBOHYD 91 91
 FT CARBOHYD 131 131
 FT CARBOHYD 262 262
 FT CARBOHYD 288 288
 FT CARBOHYD 401 401
 FT CARBOHYD 447 447
 FT CARBOHYD 469 469

RESULT 3
 CASR_BOVIN

FT CARBOHYD 489 489 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 542 542 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 595 595 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 1085 AA; 121170 MW; 5D66B8C9CD13EF7 CRC64;

Query Match 24.5%; Score 1087.5; DB 1; Length 1085;
 Best Local Similarity 31.6%; Pred. No. 9e-73;
 Matches 279; Conservative 161; Mismatches 347; Indels 97; Gaps 28;

11 LEFLVLAEPANSDPFLPGDVLGLGLSLHNMKGIVHNLQVPR---MCKEYEVKVI 67
 11 LAFSTCTSAVGPDPQAKKKGGDILLGLLEPIHGV-AVADQDKSPRESVECTRYNR-- 67
 68 GYNLMQAMFAVEIINNDSILPGVLGLGEIYDVVC-YISNNQVLYFLA-HEDNLPILQ 125
 68 GFMWQAMFAFEINSSRALLPNNMLGRIPRTCTVSKALEATSTFAQKIDSLND 127
 126 E--DYSNYSIRVAVIGPNSSEVMTVANFLSLFLPQITYSASIDELNDKVPFALLRT 183
 128 EFCNCSEHIPSTIAVVGAGSGISTAVANLGLFYIPQVSASSRLLSNKQKFSFLT 187
 184 TFSADHVMAMQMLHEFMMNITIVSSDTYGRDNGQLGEPVARDICIAQETLPL 243
 188 IPDEHQATAMADILEYEFMMNVGTIAADDDYGRPGIEKFEFEAEERDIDISELI--- 244
 244 QPNQMTSEEROLVTIYDKLOOSTARVVVSPDLTYHFENEVRONTGAVWIASHS 303
 245 ----SQSYSE-EKIQGVVEVIONSTAKVIVVSSGPDLEPLKEIYRNNITGTIMLASER 299
 304 WAIIDPVNLHMLTELGH-LGTFLLGTTIQSVPIPGFSEF-----REMGAQA-- 345
 300 MASSSLI-AMPEYFHVVGCTIGFGLKAGQIPGRREFLQVHPRKSVHNGFAKFEWETFN 358
 346 -----GPPPL-----SRTSOSYF-----CNOEDNCNANLSEFTIIRLSG 381
 359 CHLOEGAKPPLPDTFLRKHHEBGAGLNSPFAFRPLCTGE-ENISSVETPYMDYTHL-- 415
 382 ERVVSYSVAVVAHAHALSHSLGC-----DKSTCTRVVYPMOLLSEIMKVFNTL-L 432
 416 -RSTVYVLAIVSIAHALDITCIPGRGLFNGSCADIKKYAMQVLYKHLRLNTSMN 474
 433 DHQIFDPQGDVALHEIYQOMQNSQNP--FQSVASYPL-----OROLKNIQDISMHT 485
 475 GROYTFDEGDLAGNYSIINMHLSPEDGSIYFEKV-GYNNVYAKKGERLFINDEKILMSG 533
 486 VNNTIMSCSKRCGSGOKKKPV-GIHVCCFECIDLPGLFNLHNTDEHECCACPNMNS 544
 534 FSEYVFSNCSKDCLAGTKKGLIEGPTCCFECVECPDEYSDET-DASACDKCPDDEWS 592
 545 YQSETSCFRQVLEFMHAPTIYAVALLAFLSTLALIVFMRHFOPTIVSAGGPMC 604
 593 NENHTSCIAKEIEFLSWTEPFGALFLFAVLGIFLFAVIGVIRKRNPIYKATRELS 652
 605 FLMFLLLVAVVYVYVGPVYVSTCLCRQALPELCTTICISCIASVFOICAF--KMA 662
 653 YLLSLFLOCSSSLFEIEPDMTCRLRQPAFGISFVLCISCIILKRVLLVFEAKPI 712
 663 SREPRASYVWVRVYQGVYMAFITYLKMIVYIGMLARQSHPRDDPKTIIVSCNN 722
 713 TSEHRR---WGLINTOELFLVCTEPMQIVICAMLTAPSSRYNHEDELIIYFICHG 769
 723 YRNSLLENTSLDLILSVGFSFAYMGKELPTVYNEAKFTLSTKFFTSVSICTPMSAY 782
 770 SLMAAGELIGYCLLAALGFEFAFKSRKLPENNEAKFTFMSGLIFFIYWIS---PIPAY 826
 783 SGVLYTIVDLVTVNLALISIG---YGPCKYMLTFPERNT 822
 827 ASTYGRFVS-AVEVIAIILASFGLLACIFFNKYVYIILFRPSHT 869

RESULT 4
 CASR_MOUSE STANDARD: PRT; 1079 AA.
 ID CASR_MOUSE

AC Q90Y96; 008968; 088519; Q90Y95; Q90Z08; Q9R1D6; Q9R1Y2;
 DT 16-OCT-2001 (Rel. 40; Created)
 DT 16-OCT-2001 (Rel. 40; Last sequence update)
 DT 15-JUN-2002 (Rel. 41; Last annotation update)
 DE Extracellular calcium-sensing receptor precursor (CASR) (Parathyroid
 DE Cell calcium-sensing receptor).
 GN CASR OR GPRC2A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxId=10090;
 RN [1]
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
 RC STRAIN=C57BL/6; TISSUE=Kidney;
 RX MEDLINE=20092890; PubMed=10625662;
 RA Oda Y., Tu C.-L., Chang W., Ciummire D., Koemueves L., Mauro T.,
 RA Elias P.M., Biale D.D.;
 RT "The calcium sensing receptor and its alternatively spliced form in
 RL murine epidermal differentiation.";
 RL J. Biol. Chem. 275:1183-1190(2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM A), AND TISSUE SPECIFICITY.
 RC STRAIN=Black Swiss X 129/SV; TISSUE=Kidney;
 RX MEDLINE=20119279; PubMed=10652312;
 RA P1 M., Garner S.C., Flannery P., Spurney R.F., Quarles L.D.;
 RT "Sensing of extracellular cations in Casr-deficient osteoblasts.
 RL Evidence for a novel cation-sensing mechanism.";
 RL J. Biol. Chem. 275:3256-3263(2000).
 RN [3]
 RP SEQUENCE OF 256-600 FROM N.A. (ISOFORM A), AND TISSUE SPECIFICITY.
 RC TISSUE=Epiphyseal cartilage;
 RX MEDLINE=20043955; PubMed=10579354;
 RA Chang W., Tu C., Chen T.-H., Komuves L., Oda Y., Pratt S.A.,
 RA Miller S., Shoback D.;
 RT "Expression and signal transduction of calcium-sensing receptors in
 RL cartilage and bone.";
 RL Endocrinology 140:5883-5893(1999).
 RN [4]
 RP SEQUENCE OF 507-582 FROM N.A. (ISOFORM A).
 RC STRAIN=MMR; TISSUE=Brain;
 RA Hildebrand J., Ammon H.P.T., Wahl M.A.;
 RL Submitted (May-1998) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 562-814 FROM N.A.
 RC TISSUE=Kidney;
 RA Mowad T.T., Ricciardi D.;
 RL Submitted (May-1999) to the EMBL/Genbank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 646-799 FROM N.A., AND TISSUE SPECIFICITY.
 RX MEDLINE=97231187; PubMed=9076582;
 RA Quarles L.D., Hartle J.E. II, Siddhanti S.R., Guo R., Hinson T.K.;
 RT "A distinct cation-sensing mechanism in MC3T3-E1 osteoblasts
 RT functionally related to the calcium receptor.";
 RL J. Bone Miner. Res. 12:393-402(1997).
 CC -1- FUNCTION: SENSES CHANGES IN THE EXTRACELLULAR CONCENTRATION OF
 CC CALCIUM IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A (SHOWN HERE) AND B; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -1- TISSUE SPECIFICITY: EPIDERMIS, KIDNEY AND CARTILAGE.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC EMBL; AF110178; AAD28371.1; -

QY	117	-HEDNLLPIQF--DYSNTISRYVAIVIGPDNSESVMTVANFISLFLPLQITYSAISDELKD	173
DR	EMBL:	AF110179.1; AAD28372.1; -	
DR	EMBL:	AF112884.2; AAD0638.1; 1; -	
DR	EMBL:	AF068900; AAC19388.1; -	
DR	EMBL:	AF027140; BAA77688.1; -	
DR	EMBL:	AF002015; AAC53252.1; -	
DR	EMBL:	AF159565; AAF00193.1; -	
DR	MCD:	MGI:1351351; GPCR24.	
DR	InterPro:	IPR001828; ANF_Receptor.	
DR	InterPro:	IPR000337; GPCR_Mgr.	
DR	Pfam:	PF00003; Tm3_3; 1.	
DR	Pfam:	PF01094; ANF_Receptor; 1.	
DR	PRINTS:	PRO0248; GPCRMR.	
DR	PROSITE:	PS00979; G_PROTEIN_RECPE_F3_1; 1.	
DR	PROSITE:	PS00980; G_PROTEIN_RECPE_F3_2; 1.	
DR	PROSITE:	PS00981; G_PROTEIN_RECPE_F3_3; 1.	
DR	PROSITE:	PS0259; G_PROTEIN_RECPE_F3_4; 1.	
KW	G-protein coupled receptor; Transmembrane; Glycoprotein; Signal; Alternative splicing.		
FT	SIGNAL	1	POTENTIAL.
FT	CHAIN	20 1079	EXTRACELLULAR CALCIUM-SENSING RECEPTOR.
FT	DOMAIN	20 612	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	633 635	I (POTENTIAL).
FT	DOMAIN	636 649	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	650 670	II (POTENTIAL).
FT	DOMAIN	671 681	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	682 700	III (POTENTIAL).
FT	DOMAIN	701 724	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	725 745	IV (POTENTIAL).
FT	DOMAIN	746 769	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	770 792	V (POTENTIAL).
FT	DOMAIN	793 805	CYTOPLASMIC (POTENTIAL).
FT	TRANSMEM	806 828	VI (POTENTIAL).
FT	DOMAIN	829 836	EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	837 862	VII (POTENTIAL).
FT	DOMAIN	863 1079	CYTOPLASMIC (POTENTIAL).
FT	CARBOHYD	90 130	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	130 261	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	261 287	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	287 386	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	386 446	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	446 468	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	468 488	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	488 541	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	541 594	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD	594 537	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	VARSPLIC	461 537	MISSING (IN ISOFORM B).
FT	CONFLICT	45 45	A -> S (IN REF. 2).
FT	CONFLICT	304 304	L -> P (IN REF. 3).
FT	CONFLICT	410 410	G -> D (IN REF. 2 AND 3).
FT	CONFLICT	566 566	V -> A (IN REF. 2).
FT	CONFLICT	595 595	E -> H (IN REF. 2, 3 AND 5).
FT	CONFLICT	610 610	E -> V (IN REF. 5).
FT	CONFLICT	814 814	F -> L (IN REF. 5).
FT	CONFLICT	889 889	L -> I (IN REF. 2).
FT	CONFLICT	906 909	TGSN -> SGWT (IN REF. 2).
FT	CONFLICT	1057 1057	V -> M (IN REF. 2).
FT	CONFLICT	1064 1064	V -> A (IN REF. 2).
FT	CONFLICT	1076 1076	I -> V (IN REF. 2).
SO	SEQUENCE	1079 AA; 120839 MM; AAFBD8D472736D6E CRC64;	
QY	Query Match	24.4%; Score 1085.5; DB 1; Length 1079;	
	Best Local Similarity	31.8%; Pred. No. 1.3e-72;	
	Matches 285; Conservative 162; Mismatches 329; Indels 121; Gaps 31.		
DB	9	ALALATY-----HSSAAGPDQRAOKKDDIIIGLGFPIHF--GVAARD--QDIKSREE	56
QY	63	EKKVIGYNL-----MQAMFAVEEINNDSSILPGVLLGEIYDVC-YISNNQPYLYFLA	116
DB	57	SECGIRYNGRGRRWLOAMIFALEEINSSPALLPNMTLIGRIDTCTNYSKALEATLISFYA	116
QY	117	-HEDNLLPIQF--DYSNTISRYVAIVIGPDNSESVMTVANFISLFLPLQITYSAISDELKD	173

Db	117	ONKIDSLIDFENCSNHEIPTIAVNGATGCGSVSTANANLGLFIPQVAYASSRLSN	176
Qy	174	KVREPALRTTPPSADHVEANVOLMLHRENMWITVYSSDPTYGDNQOLLGERVARDIC	233
Db	177	KNOFSEFLRTIPDNEHQTAMADIEEFYRMWVSTIADDDYGRPGIEKREAEERDIDC	236
Qy	234	IAFOETLTLPONNMNTSEESORLVTIYDKLOOSTARVYVAFSDLLYHFNEVLKONF	293
Db	237	IDFESLTI-----SOYSDEEFICQVVEYI-----QNSTAKIYVSSGPDLEIKETVRNI	268
Qy	294	TGAVWIASSEMAIDPVLNLTLELGH-LGTFLGITIQVSPIPGSEF-----	338
Db	289	TGRITIALEMAASSLI-AMPEYFHVVGATIGTFLKAGIDPGEFEFLQKVPKRYVHNGF	347
Qy	339	-REMPQPA-----GPPRLSRTQSYQYCNQEDCNCAITSFMTIIRLGE-----	382
Db	348	AKEEFEELFNCHLDGAKGLPVDTFYVRSH--EEGGRNLLNLSYAFRPL--CYGDEINIS	403
Qy	383	-----RVVYSVYSAVAAVAHLSLAC-----DKSTCYKRVVYPMQOLLE	423
Db	404	VETPRMGVEHRISTVNYLAYYSIAHMLDITLCPORGFLTNSCADIKKVEAMQYLKH	463
Qy	424	IMKVNFT-LLDHQIFEPDQGVALLHLEIVQMQRSONP--FQSVASYFLP-----OROL	475
Db	464	LRLHLEFTNNMBOYVTEFDCGLDVNTSLIINHLSPEDGSIVKFEV--GYVNYAKGERLF	522
Qy	476	KNIODISMTVNNNTIPMSCKRCSROSGOKKPV-GIHVCEPCIDCJPGFTLNTDEYE	534
Db	523	INEGILMSGSRVREYFNSCRDQAGRKIIIGEPICCEYECVCDGSEGT-DASA	561
Qy	535	COACPNNMWSYQSETSCFKROLVLEHWEADTAVALLAIGLITSLAIIIVFMRHFOFP	564
Db	582	CDKCGDDPDWSENENTYSCIAKEIEFLAEMTEPRGIALTFAYLGIFLFAVLGVFIKFRNP	641
Qy	595	IYRSAGGCMCLMTLILVAVMVYVYVYKPVSTCROALFPLCTIGTICISCAVASFQ	654
Db	642	IYKANRELSTULLFSLCCSSSLFPIGEPQDMTCRLRQPAFISVYLICISLIVKTNR	701
Qy	655	IYCAF--KMASRFPRAYSWYRYOGPYVSNAFITVLMKYIVVIGMLARPOSHPTDDBP	712
Db	702	VILVEAKIPIPSFHRK--WMGLNLQFLVPLCFWQIVICIIILVYAPSSVYRHELED	758
Qy	713	KITIYSCNPNRNSLTLENTSL---DLISVYGFSPAFMKGKLEPNVYAEKFIPLSMFYF	769
Db	759	ELIETTC--HEGSLMALGSLIGYTCLIAACEFAPKRSKRLPENFEAKFIYFSMLIF	815
Qy	770	TSSVSLCTCFPMASGVGLVTIYDLIVYVNLNLAISLG---YFGCKCMILFPERNT	822
Db	816	IYWIS---FIRAYASTYCKPFS-AVEYAIALLAASFGLLACIFFNKYVIIILFKPSRNP	868
RESULT 5			
AC	MGR7_RAT	STANDARD;	PRT; 915 AA.
AD	P35400;		
DT	01-JUN-1994 (Rel. 29, Created)		
DT	01-JUN-1994 (Rel. 29, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Metabotropic glutamate receptor 7 precursor (mgIur7).		
GN	GRM7 OR GPRCIG OR MGUR7.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Sprague-Dawley; TISSUE=Brain;		
RX	MEDLINE=94117433; PubMed=8288585;		
RA	Okamoto N., Hori S., Akazawa C., Hayashi Y., Shigemoto R.,		
RA	Mizuno N., Nakanishi S.;		
RT	Molecular characterization of a new metabotropic glutamate receptor		
RT	mgIur7 coupled to inhibitory cyclic AMP signal transduction.?		

Db	71	K-RENGTHRLBMLYALDQINSDBNLLBNPTLGRILDTG--SBDYTALESQSLTFVQAL-	126
Qy	123	PIQEDYSNV-----ISRVAVITGPDNSESVMTANFLSLFLPQIYSAISDE	170
Db	127	-IQKDTSDVRCNGEPPEVFAKPEKVVYIGASGSSVSIMVANTILRLPQIDISVASTAPE	185
Qy	171	LQDKRFALLRTTPSADHNEAVYQMLNHRMMIIVYVSSDTYGGDNQGLGERAVR	230
Db	186	LSDDRRYFFSRVYVPPDSEFOAAWDYKALGMNVYSTLASSESYG-EKGVESFTQISKE	244
Qy	231	-DIDIAFOETPLPQPNQNTSEERQLY-----TIYDKLQOQSTARVYVFPDLTYHF	284
Db	245	AGGLCIASQVNP-----QERKQRTIDFDRIHQOLDTPRSRAVYFANDEDIKQI	295
Qy	285	FNEVLKRNFTG-AWAIASESA--IDPYLNLNTELGLHGLTFGLTIO--SVLPDGESEFR	339
Db	296	IAAAKRAQOVGHFHWGSDSGSKINP-LHQHEDIAE---GALTIDPKRAVYEGFPAYF	350
Qy	340	EMGPGAGPPLSRTPSSQSTCN-----QECDNC--LMTLSPTLRLISG--ERY----	384
Db	351	-----ISRLIENRRRWVVAEYEEFNKLIISGKKEDTDPRKCTGQGERIKDS	400
Qy	385	-----VYSVYSAVYAVAHALHSLLGCKSTGCR-RVYVP-----WOLLEETKRVNF	429
Db	401	NYEDEGKQVQFVIDAVYAAHALLHM---NKDLCADYRGVCEPEMBAQAGKKLTKYIRHVF	457
Qy	430	T-LDHOFTFPDQGVNLAHLEIYQWMDRSONPFOASAYTPILORLKNIQDISMHTVNN	488
Db	458	NGSACTPWPMEKNGDAGRGYDIFDYQTTNTNPGRILIGQWTDLOL-NIEDMOWGKVR	516
Qy	489	TIPMSCKRCSQSGQKKRPVGIHWCCECIDCLPGETLNMTEDEYECQACRPNNESWYQSE	548
Db	517	EIPSSVCTLPCKPGQKKTKQGTGQWMTCEPC---DDGYQVQFDMTCQHQCPYDQRPENR	573
Qy	549	TSCFRQLVYLEMEHAPRIANALLAAGFSTLALVIFMRHFTPIYVRSAGGEMCTLM	608
Db	574	TGCQNPILPIKLEHSPMAVIVPFLAMGLIATITVMAETIKYNDTPIVRASGRELSTVLL	633
Qy	609	TLLAVVWVPPVYQPPRVSTCLRQALPFLCTFICISCIIV-----RSFOICAFKMAS	663
Db	634	TGIFLCYIITFLMLAKPRVAACSRRAVPLG--GMCISYALLKTKNRIIRIPQGGKSV	691
Qy	664	REPRAVSWRYQGPVSMATIV-LKNVIYVIGMLARPOSHPTDPPDKITI---V	717
Db	692	TAPRLISPTOLA--ITSSLSIVOLLGVFIWFGV-----DPPNIIIDYDEHK	736
Qy	718	SONPWRNSLLENFSLD-LLSVYGFSS-----FAYMGKELPTVYNKAFITLSMTF	767
Db	737	TMNPQANGVLKCDITDQIICISGYSILLMWTCTVYAIKTRGVENNEKPRIGFTM--	794
Qy	768	YFTSSVSLCTFMSAYSGLVATVYD-----LVTVLNLTA-ISLG-YFGRFCYMLITYPE	819
Db	795	YTTCTIVMLAFIPIEFGTAQSAEKLYIQDTTTLTISMNLSASVALGMLYMPRVYIIRPHE	853
Qy	820	RNT---PAYFNSMIOGYTM	835
Db	854	LNVOKRKRSEFKAVVTAATM	872
RESULT 6			
MGR4-RAT			
AC	P31423:	STANDARD:	PRT: 912 AA.
DT	01-JUL-1993 (Rel. 26, Created)		
DT	01-JUL-1993 (Rel. 26, Last sequence update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Metabotropic glutamate receptor 4 precursor (mglur4).		
GN	GRM4 OR GPRC1D OR MGLUR4.		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus;		
OX	NCBI_TaxID=10116;		
RN	[11]		

RP SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RX MEDLINE=92110002; PubMed=1309649;
 RA Tanabe Y., Masu M., Ishii T., Shigemoto R., Nakanishi S.;
 RT "A family of metabotropic glutamate receptors.";
 RL Neuron 8:169-179(1992).
 RN (2)
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RX MEDLINE=93332699; PubMed=8338667;
 RA O'Hara P.J., Shepard P.O., Thoenes H., Venezia D., Haldeman B.A.,
 RT "The 11and-binding domain in metabotropic glutamate receptors is
 related to bacterial periplasmic binding proteins.";
 RL Neuron 11:41-52(1993).
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
 ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: IS WIDELY DISTRIBUTED IN THE CNS. PREDOMINANT
 CC EXPRESSION IS SEEN IN THE GRANULE CELLS OF THE CEREBELLUM.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLR6.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: M92077; -; NOT_ANNOTATED_CDS.
 DR EMBL: M90518; AAA93190.1; -;
 DR PIR: JH0563; JH0563.
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00979; G_PROTEIN_RECP_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECP_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECP_F3_3; 1.
 DR PROSITE: PS0259; G_PROTEIN_RECP_F3_4; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Multigene family.
 FT SIGNAL 1 32
 FT CHAIN 33 912
 FT DOMAIN 33 587
 FT TRANSMEM 588 610
 FT DOMAIN 611 624
 FT TRANSMEM 625 645
 FT DOMAIN 646 656
 FT TRANSMEM 657 675
 FT DOMAIN 676 699
 FT TRANSMEM 700 720
 FT DOMAIN 721 750
 FT TRANSMEM 751 772
 FT DOMAIN 773 785
 FT TRANSMEM 786 808
 FT DOMAIN 809 821
 FT TRANSMEM 822 847
 FT DOMAIN 848 912
 FT CARBOHYD 98 98
 FT CARBOHYD 301 301
 FT CARBOHYD 454 454
 FT CARBOHYD 484 484
 FT CARBOHYD 569 569
 FT CONFLICT 124 124
 SQ SEQUENCE 912 AA; 101818 MW; 336430EF19B4B577 CRC64;

Query Match 16.2%; Score 721.5; DB 1; Length 912;
 Best Local Similarity 26.5%; Pred. No. 1e-45;

Matches 240; Conservative 161; Mismatches 349; Indels 155; Gaps 37;
 QY 8 ICSLEFLL--WV---LAEPAN---SDFYLPGLDGLSLANMGMVHNLFLQYPMC 59
 Db 16 LCLLSLTPAWVSSSLKPKGPHMNSIRIDGDTTLGLLPVHGRGS-----EKKAC 67
 QY 60 KEVEKVIQYNLMQAMFAVEEINDDSSLLPGVLLGEIVDVCYISNNVQVLYFLAHD 119
 Db 68 GELK-KEKGIRLEAMFALDRINDDPLPNTIGARIDTC--SDRTALRQSLTFVQ 124
 QY 120 NLPLIQDYSNY-----ISRVAVIGPDNSESMTVAFLSLFLQIQTSAI 167
 Db 125 AL--TEKDGTEVRKSGSGPPIITKPERKVGIVASGSSVIMANILRKIPQISTAST 182
 QY 168 SDELKQKRPALLRTTPASDHHEAVQMLHFRNMVILVSDYTG-----216
 Db 183 APDLSDSRDFSRVSPDTQAQANVDYRLAKMYVSTLASSESGSGVGAFLQKS 242
 QY 217 RDNGQLGERVARRDICIAFOETPLTLPQPNOMNTSEERKRLVTVDKQSTARVVVFS 276
 Db 243 RENQ-----GVCIAQSVKIPR-EPRTGEEDKIKRL-----LETSMAGIITFA 285
 QY 277 PDLTYHFENEVLKQNTG-AVMTASENAID--PVHNLTELGHTFLGIT--QSPV 331
 Db 286 NEDDIRVLEAARRANQTHFEMWSDSGSKAPVL-RLEEVAE---GAVTLPRMS 340
 QY 332 IPGFSEFRMGPOAGPPLSRT-----SQSYCNQECNCLNATLSFTYTLRLSG 381
 Db 341 VRGDRDFS-----SRITDNNRNIMPFEMFEDNTHCKLSRRLAKSGHKKCTN 390
 QY 382 -ERY-----VYSVSAVYVAHAHLSLGCDSKSTCTRY-----VYPMQL 421
 Db 391 RERIGODSAVEQKQVFDVAVYAMGHALHAM---HRDLPGVAVGLCPMDPDVDTQL 447
 QY 422 EEIKWVNT-LLDHQIFPDGDVALLEIVOMQMSQNPQSVASYXP-LQRLQNKIQ 479
 Db 448 KYIRNVNFSGIAGNPVTFENENGDAPEGKIDYQYOLRNGSAEYKIVGWTHTLHRIEMQ 507
 QY 480 DISWHTVNNTPISMSCSKRCOSGOKKPKVGHVOCCEPIDCLPDTFLNHTDEYECACP 539
 Db 508 ---WPGSQQLPRSGISGLPCQPGERKKTVMGACMCCEPC---TGVOYQVDRYCKTCP 561
 QY 540 NNEKSVSESCRRQQLVEFEMHEAPITVAALLAAGFLSTLALVTFMHRFQPTVRS 599
 Db 562 YDKMPENRTSCQPIPIVKKLEMDSPMAVLDPLFAVVGIAATLVVVFVYVNDTPYKAS 621
 QY 600 GCPMCFLLTLIVAYVWVYVYVGPVKVSTLCRQALPFCFTICISCIARVSQIYCA 659
 Db 622 GREISYVLGIFLCVATFTLMAEPDLGCSLRIFLGLGMSIYVALLTKNRIYRIF 681
 QY 660 KMASRPPRAVSVYVRQGP--VYSMAETIV-LKMYIVIGMLARPOSHPRTPDDPKRTI 716
 Db 682 EOGKRSVSA---PRFTSPASOLAITFILLISLQIGICVWFVDP--SHSVVDQDQR-- 733
 QY 717 VSCNPNNRNLNLNTSLD-LLSVGRS-----FAFMGKELPNVNEARFYTISMT 766
 Db 734 -TLDPFRARGVLCODISLGLGYSMLMVTCTVYAKRTGVPTEFGEARPIGFTM- 791
 QY 767 FYFTSSVSLCTFMSAYSG-----VLTIVDILLVYVNLTAISLG-YFQPKCYMILFY 818
 Db 792 --YTTCTVLAFLPIPFPGTSQSDAKLYITQTTLVSVLSASVSLGMLYAPKYYIIILF 849
 QY 819 ERNTP 823
 Db 850 EQNVP 854
 RESULT 7
 MGR7_HUMAN
 ID MGR7_HUMAN STANDARD; PRT; 915 AA.
 AC 014831;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)

15-JUN-2002 (Ref. 41, last annotation update)
 Metabotropic glutamate receptor 7 precursor (mglur7).
 GNR7 OR GPRC1G OR MGLUR7.
 Homo sapiens (Human).
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NCBI_TaxID=9606;
 [1]
 SEQUENCE FROM N.A.
 TISSUE=Brain;
 MEDLINE=96437220; PubMed=8840028;
 Makoff A., Pilling C., Harrington K., Emson P.;
 "Human metabotropic glutamate receptor type 7: molecular cloning and
 RT mRNA distribution in the CNS";
 RL Brain Res. Mol. Brain Res. 40:165-170(1996).
 [2]
 SEQUENCE FROM N.A.
 MEDLINE=96141892; PubMed=9473604;
 Wu S., Wright R.A., Rockey P.K., Burgett S.G., Arnold J.S.,
 RA Rostock P.R., Jr., Johnson B.G., Schoepf D.D., Belagaje R.M.;
 "Group III human metabotropic glutamate receptors 4, 7 and 8:
 RT molecular cloning, functional expression, and comparison of
 RL pharmacological properties in RGT cells";
 RL Brain Res. Mol. Brain Res. 53:88-97(1998).
 [3]
 VARIANT PHE-433.
 MEDLINE=21095249; PubMed=11163549;
 RA Boloma A.A., Kerwin R.W., Munro J., Arranz M.J., Makoff A.J.;
 "Polymorphisms in the genes for mglur types 7 and 8: association
 RT studies with schizophrenia";
 RL Schizophr. Res. 47:99-103(2001).
 -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
 CC ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN MANY AREAS OF THE BRAIN,
 CC ESPECIALLY IN THE CEREBRAL CORTEX, HIPPOCAMPUS, AND CEREBELLUM.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLUR4
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC -----
 DR EMBL: X94552; CA64245.1; -
 DR EMBL: U92458; AAB51763.1; -
 DR Genew: HGNC:4599; GRM7.
 DR MIM: 604101; -
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00979; G_PROTEIN_RECPT_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECPT_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECPT_F3_3; 1.
 DR PROSITE: PS50259; G_PROTEIN_RECPT_F3_4; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Multigene family; Olfaction; Polymorphism.
 FT SIGNAL 1 32
 FT CHAIN 33 915
 FT DOMAIN 33 590 METABOTROPIC GLUTAMATE RECEPTOR 7.
 FT TRANSSEM 591 615 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 616 627 I (POTENTIAL).
 FT TRANSSEM 628 648 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 649 654 II (POTENTIAL).
 FT TRANSSEM 655 675 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 676 702 III (POTENTIAL).
 FT TRANSSEM 703 723 IV (POTENTIAL).
 FT TRANSSEM 723 723

FT DOMAIN 724 753 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 754 775 V (POTENTIAL).
 FT DOMAIN 776 788 CYTOPLASMIC (POTENTIAL).
 FT TRANSSEM 789 810 VI (POTENTIAL).
 FT DOMAIN 811 825 EXTRACELLULAR (POTENTIAL).
 FT TRANSSEM 826 850 VII (POTENTIAL).
 FT DOMAIN 851 915 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 98 98 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 458 458 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 486 486 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 572 572 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARIANT 433 433 Y->F.
 FT SEQUENCE 915 AA: 102250 MW: CFP94E0BFF4919 CRC64;
 SO
 Query Match 16.2%; Score 721.5; DB 1; Length 915;
 Best Local Similarity 27.9%; Pred. No. 1e-45;
 Matches 256; Conservative 152; Mismatches 354; Indels 157; Gaps 41;
 9 CSLFLLWVLAEPANSDPFLP-----GDYLLGGLFSLHNMKGIHLNFIQVCKEY 62
 19 CVLEVLICALAAAGQEMTPAHSIRIGDYLGLGFLPVAH--KGP-----SGVF-CGDI 70
 63 EKVIVGYMLQMRPAVEEINNDSSILPGVLYGEIVPVCISSNVQVLYFLAHEDNLL 122
 71 K-RENGIRHLEMLYALQINSDPMLRPVTLGARILDTG--SRDTYALDGLSTVQAL- 126
 123 PLEDYSNY-----ISRYAVIIGPNSSVMTANFLSLFLPDTYSASIDE 170
 127 -IQKTSVRCNGEPFVFKREKVGVYAGSGSSVSMVANITLFLQIPQISTASTAPE 185
 171 LDRKRFALLRTPSSADHNHEAMQMLHPRMNIIVLYSDTYGRNGQLGSRVAR 230
 186 LSDRRYDFFSKVPDPDFQADAMVDYKALGMVNSTLASGSGY-EGVESFTQISKE 244
 231 --DIIAFOETLPIQPNQNMNTSEEROLY---IYVDKLOOSTARVVVSPDLTLHYF 284
 245 AGGLCIASVRLP-----QERKDTIDEDRIIKOLDIPNRNAVIVFANDDINOI 295
 285 FNEVLKQNTG-AWVIASESNA--IDPVLHNTLGLHGTFLGTLTIQ--SVIPGFSER 339
 296 LAARARAOVGHFWLWVGSWSKINP-LHQHEDIAE--GAITYOKRAVVEGDFAYF 350
 340 EMGPAGPPLISRTSOSTCN-----QECNOC--LNATLSFTLRLSG-ERV----- 384
 351 -----TSRTLENNRRNWRFAEYWEENFNCKLTISGSKKEPDKCTGQERIGDS 400
 385 -----YYSYSAVYAVALHLSLGCDSCTK-RVYVP-----WOLLEELMKYNF 429
 401 NTEQSGKVQFVIDAYAAHAAHLMH--NKDLCADYRGVCPMEAGGSKILKTYIRNVF 457
 430 T-LLDHQIFPDQGVVALHLEIVQWQDRSONPFOVASVYPLQRLQNTIDISHTVNN 488
 458 NSAGTAPVWFNKGNDAPGRYDFIYQYTNNTSNPGRYLIGOMVDEQL--NIEDMONGKVR 516
 489 TTPMCSKRSQSGOKKRPVGIHNCFCFICDLPFTPLNHTPDEEQACAPNNKESYSE 548
 517 EIPASVCTLPCKPGQKRTQKGTGCPCCWCEP---DGYYQYDENTCCHCPDQRPNEWR 573
 549 TSCFROVLEFHEAPTAVALAALGFLSLALIVLFMRFQPIYRSAGPACFTML 608
 574 TGGCQIPILKLEHSPMAVIVPFLAMGLIINIFPMAFIRKINQPIYRASRELSYVL 633
 609 TLLVAVVAVPVYVGPVRYSTCLQALPFLCTICISCIAY-----RSQIVCAFKNAS 663
 634 TGIFLCYITTEFLMKPDVAACSEFRVFLGT--GMCISYALALTKTNIRYRFEQGRKSV 691
 664 RPRRAVSWRKQGYVEMAFITV-LKMYIVIGLARPQSHRRPDPDKITI-----V 717
 692 TAPRLISPTSLA--ITSSLSVQLDGVFTWGV-----DPPNIIIDYDEHK 736
 718 SCNPYRSLFNTSLD-LLSVGVFS-----FAYMGKELPTYNNAEFTLSMTF 767

Db 737 TMNPBARGVLCDDITDLOITCSLGLMTCTVYAIKRGVPEPNEAKPIGFTM-- 794
 QY 768 YFTSSVSLCTFMSAGSVLTIVDL-----LVTVLNLA-ISLG-YFGPKCYMLFPE 819
 Db 795 -YTTCTVWLAFIPFPGTQASAKIYIQTTLTILSNLSVALGMLKMYKVIITIFHPE 853
 QY 820 RNF---PAYFNSMIQGYTM 835
 Db 854 LNVQKRKRSEKAVVATATM 872
 RESULT 8
 MGR3_RAT
 ID MGR3_RAT STANDARD: PRT: 879 AA.
 AC P31422;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Metabotropic glutamate receptor 3 precursor (mgIur3).
 GN GNM3 OR GPRC1C OR MGUR3.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID:10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA MEDLINE=92110002; PubMed=1309649;
 RX Tanabe Y., Masu M., Ishii T., Shigemoto R., Nakanishi S.;
 RT "A family of metabotropic glutamate receptors."
 RL Neuron 8:169-179(1992).
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLYLATE CYCLASE
 ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: IS WIDELY DISTRIBUTED IN THE CNS. PREDOMINANT
 EXPRESSION IS SEEN IN THE NEURONAL CELLS OF THE CEREBRAL CORTEX,
 DENTATE GYRUS, AND GIGAL CELLS THROUGHOUT BRAIN REGIONS.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGUR2.
 CC
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M92076; -; NOT_ANNOTATED_CDS.
 DR PIR: JH0562; JH0562.
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRNGR.
 DR PROSITE: PS00979; G_PROTEIN_RECP_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECP_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECP_F3_3; 1.
 DR PROSITE: PS50259; G_PROTEIN_RECP_F3_4; 1.
 KM G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KM Multigene family.
 FT SIGNAL 1 22
 FT CHAIN 23 879
 FT DOMAIN 23 879 METABOTROPIC GLUTAMATE RECEPTOR 3.
 FT TRANSMEM 576 576
 FT DOMAIN 577 599 I (POTENTIAL).
 FT TRANSMEM 600 613 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 614 634 II (POTENTIAL).
 FT DOMAIN 635 645 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 646 664 II (POTENTIAL).
 FT TRANSMEM 665 688 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 689 709 IV (POTENTIAL).
 FT DOMAIN 710 734 EXTRACELLULAR (POTENTIAL).

FT TRANSMEM 735 756 V (POTENTIAL).
 FT DOMAIN 757 769 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 770 792 VI (POTENTIAL).
 FT DOMAIN 793 802 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 803 828 VII (POTENTIAL).
 FT DOMAIN 829 879 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 292 292 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 414 414 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 879 AA; 98959 MW; 3E5965DE5E6DEED CRC64;
 Query Match 16.2%; Score 718; DB 1; Length 879;
 Best Local Similarity 26.1%; Pred. No. 1.8e-45;
 Matches 225; Conservative 147; Mismatches 357; Indels 134; Gaps 29;
 QY 29 LPDDYLLGGLFSLHANMKGIVHNLQVPRCKEYKVGIGYKLNLMQARFAVEEINDSSL 88
 Db 35 IEGDLVVGGLFPINEGKTGTEECGRINEDR-----GIQRLNMLFADIEDINDYL 85
 QY 89 LPGLVLLGEIYDVC-----YISNNQPVLYFLAHEDN---LDP-----IOEDYSNTISRV 135
 Db 86 LPGVKLGVHILDTCSRDTVALQESLFEVRASLTKVDEAEYCPDGSVALGE---NIPILI 142
 QY 136 VAVIYGPDSSEVNTVANFLSLFLLPQITTSASIDELRDKRFPALLRTPTSDAHNRYAM 195
 Db 143 AGVIGSYSSVSIOVANLRLFLQIPISTASTAKSDKSRDYFAFATVPDPFYQAKAMA 202
 QY 196 QLMHFPMNIIIVLSDFTRGNGOLGGERVARDICIFQETLPIQNONMTSPERO 255
 Db 203 EILRFPMWTVSYVASBGDYGEGIEAFEGEALRNICITAKVGRSNIRKSYDS---- 258
 QY 256 RLVTYIDKLOOSTARVYVF---SPDLTYHFENEYLRQNTGAVWJASEMAIDPVLN 312
 Db 259 ---VIRELLQKPARVYVLEMRSDSRELIJAANRV-NASFT---WVASDGMQAQESIVK 311
 QY 313 LTELGHGTLTGLTIGSVF-----PEFSFRERPGAGRPPLSRTS 354
 Db 312 GSEHVAYGA-ITLELASHPRQDRYFQSLNPNYNNRNPWFQDFWE----- 356
 QY 355 QASYTCN-----QECDCINATLSFNTILRTSGERYVSVYSAVAVALHSLG-- 404
 Db 357 QKQCSLQGNKRNRROYCDKRLAIDS-----NYEQEKINFNNAVAYAMAHALHKQRTL 411
 QY 405 CDKST--CTRRVYVWQLL--EETWKNFTL-----LDHQIFPDQGVALLHLEIVQW 453
 Db 412 CPNTTKLCKADKILIDGKKLKEYLAKINFAPNPKNKGAISIVKFDGDMGRYVFNFL 471
 QY 454 QMRSONPQSVASVYPLQRLQKNIQDISMHTVNTIPNMSCRSKROSGKKRPVGIHV 513
 Db 472 QQTGKYSYLVKGVHW--AETLSLDVDSIHWS--RNSVPYSQSDPCAPNEMKNNQPDVC 527
 QY 514 CPECIDCLPSTFLNHEDEYECACPNNEWSYQSESCFROLVFLFEMHEAPTAVALLA 573
 Db 528 CWICICEPEPEYEL--VDEFTGMDCCGCGPPTADLSGCYNLPDXYIKWEDAMAIQPVITA 584
 QY 574 ALGFLSTLALIVFWRFQPIYRSAGGPMCFMLTLILVAVVYVYVCPVRYSTCLR 633
 Db 585 CLGFLCTCIYITFEIKNNPNLVKASGRELCYILLGVSLSYCMTEFFFIKPSPVICAR 644
 QY 634 QALFPLCTFCISCIASVQIYCAF---KMASRFPAYSYWRYQGP---VYSMAFIT 686
 Db 645 RLGLGTSFALCYAALLTKTKICARIFDGVKNGARRRFIS-----PSSQVITCIGLLI 697
 QY 687 VLKNVYVIGMLARPOSHPTDPPDKITIVSCNPYRNSLNTSLDLLSVGSEFAY 746
 Db 698 VQLVAVSWMLILEPGRRTYLPDEKRETVLKCNVK--DSMLSLAFVDVYLICVYAF 756
 QY 747 MGKELPYNINAEAFITISMTFTSSVSLCTFMSAGSVLTIVDLYL-VLMLLAISLG 805
 Db 757 KTRKCEPNEAEKPIGFTM--YTTCTIWLAFPIF---YVTSDDYRVQTTMNCISVLS 810
 QY 806 -----YFGPKCYMLFPERN 821

```

      | | | : : | | | : : |
Db 811 GFVVLGCLFAPKVHIVLFQPKN 833

```

```

RESULT 9
ID MGR4_HUMAN STANDARD; PRT; 912 AA.
AC Q14833;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metabotropic glutamate receptor 4 precursor (mglur4).
GN GRM4 OR GPRC1D OR MGLUR4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96346635; PubMed=8738157;
RA Makoff A., Leichuk R., Oxer M., Harrington K., Emson P.;
RT "Molecular characterization and localization of human metabotropic
RT glutamate receptor type 4.";
RL Brain Res. Mol. Brain Res. 37:239-248(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98141892; PubMed=9473604;
RA Wu S., Wright R.A., Rockey P.K., Burgess S.G., Arnold J.S.,
RT "Rostek P.W. Jr., Johnson B.G., Schoepf D.D., Belagale R.M.;
RT "Group III human metabotropic glutamate receptors 4, 7 and 8:
RT molecular cloning, functional expression, and comparison of
RT pharmacological properties in RGT cells.";
RL Brain Res. Mol. Brain Res. 53:88-97(1998).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=95342351; PubMed=7617140;
RA Flor P.J., Lukic S., Rueegg D., Leonhardt T., Knoepfel T., Kuhn R.;
RT "Molecular cloning, functional expression and pharmacological
RT characterization of the human metabotropic glutamate receptor type
RT 4.";
RL Neuropharmacology 34:149-155(1995).
RN [4]
RP VARIANT ILE-797.
RX MEDLINE=21416233; PubMed=11525421;
RA Ohtsuki T., Toru M., Arinami T.;
RT "Mutation screening of the metabotropic glutamate receptor mglur4
RT (GRM4) gene in patients with schizophrenia.";
RL Psychiatr. Genet. 11:79-83(2001).
CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLYL CYCLASE
CC ACTIVITY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- TISSUE SPECIFICITY: STRONGLY EXPRESSED IN THE CEREBELLUM.
CC EXPRESSED AT LOW LEVELS IN HIPPOCAMPUS, HYPOTHALAMUS AND
CC THALAMUS. NO EXPRESSION DETECTED IN LIVER.
CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST, TO MGLUR6.
CC -----
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/isb-sib.ch).
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X80818; CAAS6784.1; -
DR EMBL: U92457; AAB51762.1; -
DR Genew; HGNC:4596; GRM4.
DR MIM: 604100; -
DR Interpro: IPR001828; ANF_receptor.

```



```

Db 457 GIAGNPVTFFENNGAPGRYDIYQYOLRNDSEAEYVIGSW--TDHLHLRIRMHMPGSGQO 514
OY 490 IPMSMCKROSGOKKKPVGIHCVCEFCIDLPGETFLNHDEDEYCOACPNMNSYSOSET 549
Db 515 LPSISICSLPQOPGRKRTVGMPCMCWCERC---TGYYQYDRTYCTCPCYDMRPENRT 571
OY 550 SCRRQLVFLFEMHAPPIAVALAALGFLSTLALVLFMRHFPQPIYVSAGGPMCLMLT 609
Db 572 GCRPIPIKLEMGSPMAVLPFLFVAGIAATLFLVITFVRXNDPIYKASGRELSTYLLA 631
OY 610 LLVAVVAVVYVPPRVSCICRQALFPLCFITICICIAVRSQIYCAKMASRFRAY 669
Db 632 GFLCYATFELMAEPDLGTCISLRIFLGMSISYALLTIRIRYRIFRQGRSYSA- 690
OY 670 SYWRVQGPVYSMAF---IVLKNVIVIGLARPOSHPRTPDPDPKITIYSCPNRYS 726
Db 691 ---PRFISPASQALITSLISLQIGCVWFVDP-SHVVYEDQDQ---TIDPRARG 742
OY 727 LLENTSLDL-LLSVGFS-----FAYMKEPLPNYNEAKETILSMTEYFTSVSLC 776
Db 743 VLKCDISDLICLLGYSMLMTVCTVYAIKTRGVPELFEAKPIGFTM---YTTCIYVL 799
OY 777 TEMSAYSG-----VAVTIYDLYVTLNLAISLG-YRQPKYMIIFYPERNTP 823
Db 800 AFPIPIFGTSQASADKLIYQTTTLVSVLSASVSLGMLMPKYIILFHPQNPV 854

RESULT 10
MGR3_HUMAN STANDARD: PRT: 877 AA.
ID MGR3_HUMAN Q14832;
AC 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metabotropic glutamate receptor 3 precursor (mglur3).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=96437205; PubMed=8840013;
RA Makoff A., Volpe F., Leichuk R., Harrington K., Emson P.;
RT "Molecular characterization and localization of human metabotropic
RT glutamate receptor type 3.";
RL Brain Res. Mol. Brain Res. 40:55-63(1996).
CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLYLATE CYCLASE
CC ACTIVITY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST, TO MGLUR2.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: X77748; CAAS4796.1; -
CC Genew: HGNC:4595; GRM3.
CC MIM: 601115; -
CC InterPro: IPR001828; ANF_receptor.
CC InterPro: IPR000337; GPCR_Mgr.
CC Pfam: PF00003; 7tm_3; 1.
CC Pfam: PF01094; ANF_receptor; 1.
CC PRINTS: PR00248; GPCR_MGR.
CC PROSITE: PS00979; G_PROTEIN_RECP_F3_1; 1.

```

```

DR PROSITE: PS00980; G_PROTEIN_RECP_F3_2; 1.
DR PROSITE: PS00981; G_PROTEIN_RECP_F3_3; 1.
DR PROSITE: PS00259; G_PROTEIN_RECP_F3_4; 1.
KM G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Multigene family.
FT STGNAL 1 20
FT CHAIN 21 877
FT DOMAIN 21 574
FT TRANSMEM 575 597
FT DOMAIN 598 611
FT TRANSMEM 612 632
FT TRANSMEM 633 643
FT TRANSMEM 644 662
FT TRANSMEM 663 686
FT TRANSMEM 687 707
FT DOMAIN 708 732
FT TRANSMEM 733 754
FT TRANSMEM 755 767
FT TRANSMEM 768 790
FT TRANSMEM 791 800
FT TRANSMEM 801 826
FT DOMAIN 827 877
FT CARBOHYD 207 207
FT CARBOHYD 290 290
FT CARBOHYD 412 412
FT CARBOHYD 437 437
SQ SEQUENCE 877 AA; 98619 MW; 66F28663CE35F740 CRC64;

```

Query Match 16.0%; Score 709; DB 1; Length 877;
 Best Local Similarity 25.8%; Pred. No. 8.4e-45;
 Matches 224; Conserved 145; Mismatches 356; Indels 142; Gaps 28;

```

OY 29 LPDGLVGLGFLSHANKGIIVHLNFILOVPKCEYEVVIGYINLMQANFAVEIINDSL 88
Db 33 IEGDLVYGLFPIINEKGTGTECGRINEDR-----GIQRLMALFAIDEINKDYL 83
OY 89 LPGLVGLYEIVDC-----YISNNVQVLYFLAHEON--LLP-----IQEDSNYSIRV 135
Db 84 LPGLVGLVHLIDCSRDYVALEOSLEFVRASLTKVDAEYMCDPGSAIIDE--NIPLLI 140
OY 136 VAVIGPNSESVMTVANFLSLFLPQITYSAISDELBDKRYFPALLTFTPSADHVEAMV 195
Db 141 AGVIGGSYSVSIOVAILMLFQIPQISVASTAKLSKDRYDFATVPDPYQAKAMA 200
OY 196 QLMHFRNMNIIYVSSDYGRDNGQLGERVARIDICIAFQETLPTLPQNMNTSEERQ 255
Db 201 EILTFEWMYVSVTVASEGDEGEIEAFQEARLRNICITAEVYGRSNIRKSYDS---- 256
OY 256 RLVTIYDKIOSTARVYVSPDLTLYHFNEVILRON--FTGAVWIASSEMAIDPVILNL 313
Db 257 ---VIRELLQKPNARVYVLFMRSDSRELIAAASRANASFT---VWASDQMGQOESILKG 310
OY 314 TELGHGTFPIGITIOSVPI-----PGSEPREMPQAGPPPLSTSQ 355
Db 311 SEHYAIGA-ITLELASPVQDFRYFOSLNPYNNHRPFRDWE-----Q 355
OY 356 SYTCNOE-----CNCINATLSEFTIILRSGERYVSVYSVAVAHALSHLSLG--C 405
Db 356 KFGCSLQNKNNHRYCCKHALIDSS-----NYRDESKIMEVVAVYMAALHLKMQRTLC 410
OY 406 DKST--CTKRVPVYPMOLL--EELMKVNFLL-----LDHQIFPDQGVALLHLEYVQW 454
Db 411 PNTFKLDAMKILDKGLKLYKDYLLKINFYAPFNPKDADSIYKFDGDMGRVYFN-- 468
OY 455 WDSQNFQSVASYTPQR-----QLNKIDISMHYVNNYIPMSCKROSGOKKKPVG 509
Db 469 -----FQNVGKYSYLKGHMAETLSDVNSTIHSRNSVPTSQSCDPCAPNEMKMOP 521
OY 510 IHVCCFECIDLPGETFLNHDEDEYCOACPNMNSYSOSETSCRRQLVFLFEMHAPPIAV 569
Db 522 GDVCCWICIPCPEYEL--ADEFTCMDCSGGQMPFADLTGCTDLPEDYIRWEDAKAIGP 578
OY 570 ALTAALGFLSTLALVLFMRHFPQPIYVSAGGPMCLMLTLVAVVAVVYVPPRVST 629

```

```

Db 579 VTIACLGMCCTCMTVTVIKNNINPLVKASGELCTIILFGVGLSTCYMTFFIAPSPVYI 638
OY 630 CLCROALPFLCTTICISCIIVARSFOIYCAF---KMASFPFRAYSWYVYOGP---YVSM 682
Db 639 CALRRLGSSGSAFYCSALFTKTCIARLFDQVKNAGRAPKFIIS-----PSSGVFICL 691
OY 683 AFITVLMKAVIYVIGMLARPOSHRPDPDKTIVSCNPENRNSILNLTSLDILLSVYGF 742
Db 692 GILIVQIYVWSVWMLLEAPGRTYRLAKRETVILKCNVK-DSSMLISLTYDVILVILCT 750
OY 743 SFAYWKEPLTYNNAKFTITLSMFTYSVSLCTFMSASGVLTIVDLY-IVLNLIA 801
Db 751 VYAFETKRCPEFNENAKTIGFTM---YTCTIMAFLEPIF---YTSSDYRQVTTMCIS 804
OY 802 ISLG-----YEGPKCYMLLEFPERN 821
Db 805 VLSGFWVLVGLFAEKVHIIILFOPQKN 831

RESULT 11
MGR2_RAT STANDARD; PRT; 872 AA.
ID MGR2_RAT STANDARD; PRT; 872 AA.
AC P31421;
DE 01-JUL-1993 (Rel. 26, Created)
DE 01-JUL-1993 (Rel. 26, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Metabotropic glutamate receptor 2 precursor (mglur2).
GN GRM2 OR GPRC1B OR MGLUR2.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=92110002; PubMed=1309649;
RA Tanabe Y., Masu M., Ishii T., Shigemoto R., Nakanishi S.;
RT "A family of metabotropic glutamate receptors.";
RL Neuron 8:169-179(1992).
CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS
CC MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE ACTIVITY.
CC MAY MEDULATE SUPPRESSION OF NEUROTRANSMISSION OR MAY BE INVOLVED IN
CC SYNAPTOGENESIS OR SYNAPTIC STABILIZATION.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC EXPRESSION IS SEEN IN GOLGI CELLS OF THE CEREBELLUM AND SOME
CC PARTICULAR NEURONAL CELLS IN OTHER BRAIN REGIONS.
CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST, TO MGLUR3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC The European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M92075; -; NOT_ANNOTATED_CDS.
DR PIR; JH0561; JH0561.
DR InterPro; IPR001828; ANF_receptor.
DR InterPro; IPR000337; GPCR_Mgr.
DR Pfam; PF00003; 7tm_3; 1.
DR Pfam; PF01094; ANF_receptor; 1.
DR PRINTS; PR00248; GPCRMR.
DR PROSITE; PS00979; G_PROTEIN_RECP_F3_1; 1.
DR PROSITE; PS00980; G_PROTEIN_RECP_F3_2; 1.
DR PROSITE; PS00981; G_PROTEIN_RECP_F3_3; 1.
DR PROSITE; PS0259; G_PROTEIN_RECP_F3_4; 1.
DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW Multigene family.
FT SIGNAL 1 18 POTENTIAL.

```

```

FT CHAIN 19 872 METABOTROPIC GLUTAMATE RECEPTOR 2.
FT DOMAIN 19 567 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 568 590 I (POTENTIAL).
FT DOMAIN 591 604 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 605 625 II (POTENTIAL).
FT DOMAIN 626 636 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 637 655 III (POTENTIAL).
FT DOMAIN 656 679 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 680 700 IV (POTENTIAL).
FT DOMAIN 701 725 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 726 747 V (POTENTIAL).
FT DOMAIN 748 760 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 761 783 VI (POTENTIAL).
FT DOMAIN 784 793 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 794 819 VII (POTENTIAL).
FT DOMAIN 820 872 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 203 203 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 547 547 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 872 AA; 95773 MM; 1874CABDAD4BED9 CRC64;

Query Match 15.7%; Score 699.5; DB 1; Length 872;
Best Local Similarity 26.8%; Pred. No. 4,2e-44;
Matches 234; Conservative 157; Mismatches 365; Indels 117; Gaps 33;

OY 14 LLM-VLAEPANSDYFLPDYLLGGLSLH-----ANNKGIYHLEFLQVPMCKEYEVYI 67
Db 12 LLMGAVAGPAKKVLTLEGLDVLGFLPVHQGPAEBCGPVN-----EHR----- 57
OY 68 GYNLMQAMRPAVEEINNDSSILPGVLGLEYIVDYCYISNN-VQPIVYFLAHE-----DNL 121
Db 58 GIORLEAMLFADLRINRPHLPGLVRLGAHLIDSCSKDTHALQALDVPVRLASLGAGDS 117
OY 122 LPIQED-----YSNYSIRVAVTIGPDNSESVTVAFLSLFLPQITYSATSDHLDRVR 176
Db 118 RHICPDGSAVYAHSDAPVAVTGIVGSGSDVSIQVANLRLRQIPDISVASSAKISDKSR 177
OY 177 FPAALRTPSADHVEAMVQLMHLFRMWITVYVSSDPYGRDNGOLLGERVARDICIAF 236
Db 178 YDYFAITVPPEFFQAKAAEILRFENMTYVSIVASEGDTGELAELEAARARNICVAT 237
OY 237 QETLTLPQNMNTESEORLVTIYDKLOQSAARVYVE--SPDLTFHFENEVYRQNT 294
Db 238 SEKV-----GRMSRAAEVGVRL--LQPSARAIVFTSSEDRKELLATQRLNAST 290
OY 295 GAVMIASESW-AIDPVLHNLTELGLTFLITIOSVPIPFSEERE---WGPOAGPPP 349
Db 291 --WVASDGMGLAESVVG-SERAEGA-ITIELASYPLISDPASVYFQSLDPMNNSRNPMF 345
OY 350 LSRTSQSTCNOECNCLNATLSFTIIRLGERVYVYSVAVAVARALSL--LGCDK 407
Db 346 REFWEERHCSPRODCAHSL--RAVPFDESKIMVAVAMAHALNHRALCPN 402
OY 408 ST--C-TRVYVPMOLAE-IMKVFTL-----LDHGIFFPOGDVALHLEIYQM--OMD 456
Db 403 TTHLCDAMRPVNGRRLYDFVLANFADAPFRPADIDDEFRDRGDSGIRGINIFLYLAG 462
OY 457 RSONPFQSVASYYPQLQRLKNIQDISMTVN-NTIPMSGSRKCSGOKRRPVGIHVCF 515
Db 463 SGRYRYQKVG--YMAEGILDTDSFIPMASPAGPLPARCSGPCQNVKSVQPEVCWM 520
OY 516 ECIDCLPGFTLNHDEDEYEQACPNNEYSQSECFRQVLFLEWHBAPIYAVALLAL 575
Db 521 LCIPQOP---YEYRDEFTCADCGIGWPNASLTGCFELPDYIRWGMAMVAVPTIACL 577
OY 576 GPLSLATLIVFMRHFORPIYASAGPMLTLTLLAVYVWPVYVGPVSTICRQA 635
Db 578 GALATLFLVIGVFRHNATPVKASGRELCTYILGCVFLCYCTVFIKPPSTAVCTTLRL 637
OY 636 LFPFLCTICISCIIVARSFOIYCAFEMASRFPRAYSWYVYOGP---YVSNAFITVLMKV 691

```


Db 142 KNASSGPVGVIGGSSVSLQVANLRLRHHIPOVSPASTAKTISDKTRPDLFARTYPPD 201
 QY 188 DHRVAVQMLHFRNMWITVYSSDYTGNDGOLLGBRVARRDICIAFOETLPTLOPNO 247
 Db 202 TFOVSVALVDILKNFNMWVSTHSEGSYGEIGIALHKEATERNVCIATAKRVASADDK 261
 QY 248 NMTSEORRLVTVIKLOOS--TARVVVYFSPDLTLVHFNFVNLQNFNGAN-WTASESMA 305
 Db 262 VEDS-----ISKLOKRNAGVULFTAEARRLQAKANLSQPHWTASDGMG 313
 QY 306 -IDVHLNLTGHLGFTLGTIOSVPLPGEF-----REWGPAGPPPLSRT 353
 Db 314 KOQKLEBEDIAE--GAIVELQSEIIADPRMQLPTNQBNPFAVWEDTFNCV 371
 QY 354 SOSTCNOECNCINAT-----LSFTILLSGRVYVS-----YSAVYAVAH 397
 Db 372 LTSLSVPRDTSNSANSNDKIGVAKTECDSDYRLS--EKVGYEODESKYQVAVYAFAY 430
 QY 398 ALHSL-----LGCDKSTCRKVYVPMOLLEIN----- 425
 Db 431 ALHMLHNDRCWQSDOTTERKHLQS--ESVYRKISTYTKSQACPDMANYGKERYN 486
 QY 426 --KVNFTILD--HQIFPDQGDVALHLEIVQWQMSQNPFO--SVASYELQRLKNIQ 479
 Db 487 NYLLNVSFIDLAGSEVKEFRQDGLARYDILNYORENSSGYQKVGKWFNGIQL-NSE 545
 QY 480 DISMHTVNNITPMSCSKROSGOKKRVGHHVCCFECIDCLPFTFLNHTDEDECACF 539
 Db 546 TVVNNKETEQ-PTISACSLPCCEVGMKIKQOG-DTCWCWICDSC--ESFEYVDEFTCKDCG 600
 QY 540 NNEWSYSETSCKFROLVLEHMEAPTAVALLAALFSLTALIVFMRHPTPIVYSA 599
 Db 601 PGLMPYADKLSCYALDIQYMKMNSLFLIPMAIAIFGALTSIYVLEAKHDPVLRAS 660
 QY 600 GGPICFMTLLILVAVMYVYVGPVRYVSTCLQALFPLCFTICISCIASVQIYCAF 659
 Db 661 GRELSYLLFGLVLCYCTFALINAKPTIGSCVLOFRGIVGFSIYSLMKTNRISRF 720
 QY 660 KMASRFRAYSIVWYRGOPYSMAFITVLMKVIVIGMLARQSHPRDPDPKITIYSC 719
 Db 721 HSAKSMORLKY-ISPQSYVITTSIAIQLVITIMWVPPGRTYRPPRREY-ILKC 778
 QY 720 NPNYRNSLLENTSLDLLSVYGFSPAYMGKELPTYNENAKETILSMFYETSVSLSCTFM 779
 Db 779 KIQ-DMEFLFQSLYNMILITICTIYAKTRKIPENENSKRIGFTM--YTTCLIMLAVF 834
 QY 780 SAYSG-----VLTIVDLATVNLALISGIFGPKCYMLFPERN 821
 Db 835 PIYFGTGSYEVOTTTICISLSASVAL-VCLYSPRYIIVFHPDKN 881
 RESULT 13
 MGR2_HUMAN
 ID MGR2_HUMAN STANDARD: PRT: 872 AA.
 AC Q1416; Q9H3N6;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metabotropic glutamate receptor 2 precursor (mglur2).
 GN GRM2 OR GPRCIB OR MGLUR2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NC NCB1_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95346007; PubMed=7620613;
 RA Flor P.J., Lindauer K., Puttner I., Ruegg D., Lulik S., Knopfel T.,
 RA Ruhn R.;
 RT "Molecular cloning, functional expression and pharmacological
 RT characterization of the human metabotropic glutamate receptor type
 RT 2.";
 RL Eur. J. Neurosci. 7:622-629(1995).

RN [2]
 RP SEQUENCE FROM N.A.
 RA Yasuyuki F., Akiko J.;
 RT "Structure and polymorphisms of the human metabotropic glutamate
 RT receptor type 2 (mglur2) gene: analysis of association with
 RT schizophrenia.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS
 CC MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE ACTIVITY.
 CC MAY MEDIATE SUPPRESSION OF NEURONAL TRANSMISSION OR MAY BE INVOLVED IN
 CC SYNAPTOGENESIS OR SYNAPTIC STABILIZATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED IN DIFFERENT REGIONS OF THE
 CC ADULT BRAIN AS WELL AS IN FETAL BRAIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLUR3.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: L35318; AA76855.1; -
 DR EMBL: AB045011; BAB19817.1; -
 DR Genew: HGNC:4594; GRM2.
 DR MIM: 604099; -
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00979; G_PROTEIN_RECPE_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECPE_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECPE_F3_3; 1.
 DR PROSITE: PS50259; G_PROTEIN_RECPE_F3_4; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Multigene family.
 KW
 FT SIGNAL 1
 FT CHAIN 1
 FT DOMAIN 19
 FT TRANSSEM 568
 FT DOMAIN 591
 FT TRANSSEM 605
 FT DOMAIN 626
 FT TRANSSEM 637
 FT DOMAIN 656
 FT TRANSSEM 679
 FT DOMAIN 700
 FT TRANSSEM 701
 FT TRANSSEM 726
 FT DOMAIN 748
 FT TRANSSEM 761
 FT DOMAIN 783
 FT TRANSSEM 794
 FT DOMAIN 820
 FT CARBOHYD 203
 FT CARBOHYD 286
 FT CARBOHYD 338
 FT CARBOHYD 402
 FT CARBOHYD 547
 FT CONFLICT 12
 FT CONFLICT 210
 FT CONFLICT 496
 FT CONFLICT 496
 FT CONFLICT 748
 FT CONFLICT 776
 FT CONFLICT 776
 SQ SEQUENCE 872 AA; 95507 MW; 058608C35C701EDD CRC64;
 Query Match 15.5%; Score 686.5; DB 1; Length 872;
 Best Local Similarity 26.6%; Pred. No. 3.9e-43;
 Matches 233; Conservative 153; Mismatches 364; Indels 127; Gaps 31;

QY 15 LM-VLAEPNEDPFLPGDYLGLFSLH-----ANKGIVHLNFTLOVPCKEYEVKVI 68
 DB 13 LMGAIVEGPAKKVLTLEDGLVGLFPPVHKGKGPAAEDCGFVN-----EHR--G 58
 QY 69 YNLMQAMRFVVEEINNDSSLPGVILGYEIVDYCYISNNQPVLYFLAHDNLLPIQEDY 128
 DB 59 IORLEAMLEFALDINRDPHLLPGVRIGAHILDSK--SKDHAELOALDEVRASLSRGADG 116
 QY 129 SNYI-----SRVAVIIPDNSESVMVFANFSLFLLPQIYSAISDELDRKY 175
 DB 117 SRHICDGSATGADPAITATGYIGSYSDVSTOVANLRLPQLPQIYSTAKLSKDS 176
 QY 176 REPALLRTTPSADHVEAMQMLHFRKMTIYLVSSDIYGRDNGQLLGERVARDCICIA 235
 DB 177 RYDFYARTVPPDFEOKAKAMEILIRFEWTVYSTEASEGDEIGIEAFELAEARNICVA 236
 QY 236 FORTLPLDONMTSEBRRLTYDKLOOSTARVYVY--SPDLTYHFEHNEVLKONT 233
 DB 237 TSEKVV-----GRAMSRAAFEGVVRAL--LOKPARAVALTRESEDAELLASQRLNASF 289
 QY 294 TGAWMIASESM-AIDPVLHMLTELGHITPILG-ITIO--SVPIPGFSEPRE---MGPOA 345
 DB 290 T---WASDOWMGALLESV-----AGSSEGAELGATITIELAIPYISDPASTYQSIDPMWNSR 341
 QY 346 GPPPLRTSOSYTCNOECNCLNANTLSEFNTLRISGERVYVYSAVYVAHALHSL--L 403
 DB 342 NPWFREFWEOFRFCSEFORCDAASHL---RAVPEDESKIMFVNAVAYAMAHALHMHRA 398
 QY 404 GCDKST-----CTKRVVYPMQLLEIKVNFETLD--HOIFPDQGVALHLELYOV 453
 DB 399 LCNTRTRLCAMRPVNGRRLYKDFVLNWKDAPFRPDITNEVRFEDRFGGIGRYNIETY 458
 QY 454 -QMDRSQNPQASVASYYPLOROLKNIQDISMHTVN--NTIPMSCKRCSGOKKRPVGH 511
 DB 459 LRAGSRKRYRQYKG--YMAEGLTIDTSLIPWASPSAPPLAASRCEPCLONEKVSQPG 516
 QY 512 VCFEEICDLPGRFTLHTEDEYECQACPNMENSQSTSCFKQIYVLEHNEAPTIYAL 571
 DB 517 VCCWMLCPCP---YERIRDEFACDGLGYWPNASLTGCEIPEXIRMGDAAWGPVT 573
 QY 572 LAAIGELSTLAIIVIEFRHPOTIVRSAGGPMCFMLTLLIVAVMVVYVGPVKYSTCL 631
 DB 574 IACGLALATIFVIGVFRHNAITPVKASRELCYIILGIFLCYCMFTIARPSRACVT 633
 QY 632 CROALPPLCTTICISIAVBSFQIVCAFKNASRPRAYSTWVRQGP---YVSMATIV 687
 DB 634 LRRLIGLTAASYCYSAALTITRNIRIARIFGGA---REGAORPRFISPASQVAILALISG 689
 QY 688 LKNVIVVIGMLARPOSHPRDPPDKTITVSCNPNRYN--SLNNTSIDLLSVYGFSAF 746
 DB 690 QLLIVAVMLVEAPGTGKETAPERREVVTLRC--NHRDASMLSLAVNLLILCTLYAF 747
 QY 747 MGKELPTNYEAKFITLSM-----TFYFTSS-----VSLCTMSAVSGLVY 788
 DB 748 NTRKCEPNEAFRIGFTWTTCTIIMALLPIFYVSSDDRVQTTMCVSVLSGSV--- 804
 QY 789 IVDLTVLNLALISLUGYEGPCYMIIFPERNTPAY 825
 DB 805 -----VLGCL-----FAPKLIILIFPOKKNVSH 828
 RESULT 14
 MGR5_RAT
 ID MGR5_RAT STANDARD; PRT; 1203 AA.
 AC P31424;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metabotropic glutamate receptor 5 precursor (mglur5).
 GN GRM5 OR GPRC1E OR MGLUR5.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxId=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Brain;
 RX MEDLINE=92317054; PubMed=1320017;
 RA Abe T., Sugihara H., Nawa H., Shigemoto R., Mizuno N., Nakanishi S.;
 RT "Molecular characterization of a novel metabotropic glutamate
 RT receptor mglur5 coupled to inositol phosphate/Ca2+ signal
 RT transduction.";
 RL J. Biol. Chem. 267:13361-13368(1992).
 RN [2]
 RP SEQUENCE OF 859-923 FROM N.A., AND ALTERNATIVE SPLICING.
 RC TISSUE-Brain;
 RX MEDLINE=934343913; PubMed=7688218;
 RA Minakami R., Katsuki F., Sugiyama H.;
 RT "A variant of metabotropic glutamate receptor subtype 5: an
 RT evolutionally conserved insertion with no termination codon.";
 RL Biochem. Biophys. Res. Commun. 194:622-627(1993).
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS
 CC MEDIATED BY A G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-
 CC CALCIUM SECOND MESSENGER SYSTEM AND GENERATES A CALCIUM-ACTIVATED
 CC CHLORIDE CURRENT.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 5A (SHOWN HERE) AND 5B: ARE
 CC PRODUCED BY ALTERNATIVE SPLICING. THEY DIFFER BY AN INSERTION OF
 CC 32 RESIDUES.
 CC -1- TISSUE SPECIFICITY: WIDELY DISTRIBUTED IN NEURONAL CELLS OF THE
 CC CENTRAL NERVOUS SYSTEM.
 CC -1- MISCELLANEOUS: ACTIVATED BY OUTSOUALATE > GLUTAMATE > IBOTENATE >
 CC TRANS-1- AMINOCYCLOPENTYL-1,3-DICARBOXYLATE.
 CC -1- SIMILARITY: BELONGS TO FAMILY 5 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLUR1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: D10891; ? NOT ANNOTATED_CDS.
 CC EMBL: S64315; AAB27666.1; --
 DR PIR: A42916; A42916.
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; Tm3.1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00979; G_PROTEIN_RECPEP_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECPEP_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECPEP_F3_3; 1.
 DR PROSITE: PS00259; G_PROTEIN_RECPEP_F3_4; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Multigene family; Alternative splicing.
 FT SIGNAL 1 20
 FT CHAIN 21 1203
 FT DOMAIN 22 578
 FT TRANSMEM 579 601
 FT DOMAIN 602 615
 FT TRANSMEM 616 636
 FT DOMAIN 637 647
 FT TRANSMEM 648 666
 FT DOMAIN 667 692
 FT TRANSMEM 693 713
 FT DOMAIN 714 736
 FT TRANSMEM 737 758
 FT DOMAIN 759 771
 FT TRANSMEM 772 794
 FT DOMAIN 795 800
 FT TRANSMEM 801 826
 FT DOMAIN 827 1203
 FT CARBOHYD 88 88

FT CARBOHYD 209 209 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 377 377 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 381 381 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 444 444 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 733 733 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VASPLIC 876 907 MISSING (IN ISOFORM 5A).
 SQ SEQUENCE 1203 AA; 131885 MW; 99CA51E911C1EA4 CRC64;

Query Match 15.4%; Score 685; DB 1; Length 1203;
 Best Local Similarity 25.8%; Pred. No. 7.7e-43;
 Matches 229; Conservative 152; Mismatches 341; Indels 166; Gaps 32;

28 YLPDYLGLFSLHANKGIVLNFLOVPMCKEYEVK-----IGNIMOAFAYEE 81
 DB HMPGIIIGALFVH-----H-----QPTVDKVERKCGAVREOYGIORVAMLTHER 78
 OY INNDSSLLPGVLLGTEIVDVCYIS-----NNVQPVLYFLAHBNLLPIODEDSYNIS 133
 DB INSDPTLLPNTLIGELINDSCWHSVALEQIEFIRDSLISEEDELGYRCVDSPPRS 138
 OY 134 R--VVAIVGPNSSEVMVAVPLSLFLLPOTYSASIDELDKVAFPALRTTPADHIV 191
 DB KPIIVGIVGPSSVAIVQVNLQJLNPOLATYSATSDLSKTLFKYFMVAVPSDAQOA 198
 OY 192 EAMVQMLHFRMNIIVYSSDPTGGRDNGOLGERVARRDICIAFOETLPTLPQNMNTS 251
 DB RAMVDIVKRYMVTYSAVHTGNGESGMEAFKDKSAKEGICIAISYI-----YSN 250
 OY 252 EERQRLVIVDKLOO--STARVYVVSDDLTYHFFENYLRONTGA--VWIASSEWADP 308
 DB 251 AGESEFDLKLKRLSHLRKARVAVACFCEGMYRGLLMMARRLGLAGEFLLSGSDMA-- 307
 OY 309 VLHNLTELGHTFGITIT--OSVPI-----PGSFREMGQAPRP 348
 DB 308 DRVIDTDYQREAVAGITIKLOSPPVKFDDYILKRETNLRNPFQEFMOHRCLE 367
 OY 349 PLSTRSQSYTCNQECDCNLAFTLRLSGERYVS---VYSAYAVAHLSLIG 404
 DB 368 GFAQNSRY--NKTG-----NSSLLTLRTHVQDSKMGFVINALIYSMAVJAHNM-- 413
 OY 405 CDKSTC-----TKRYYVPMQLLEIWKVNT--LDHOFEPDQGVALLHLEIVQNO-- 454
 DB 414 -QMSLCPGAGLCDAKMDIDGRKLDSIMKNTFTGVSGDMILFDENGSPGRYEIMNKE 472
 OY 455 -----MDRSQNPQSVASYRPLQRLKIDIDISHTYNNITPMSCSKROQS 501
 DB 473 MKDYFDYINQSWDNG-----ELKMDDEVMSKKNIT--RSVCSEPEK 516
 OY 502 GQ---KKRPVGIHVCCEPCIDCLPGTFLNHTDEDECOACPNNEWSYQSESCFRQOLV 557
 DB 517 GQIKVIRKGEVS--CCWTCRCKRENEY--FDEYTCACQAGLSWPPDDLIGCLLIYQ 570
 OY 558 FLEHHEAPTIIVALLAALGFLSTALIVIFMRHFOPTIYRSAGGPMCLMTLLIVAVV 617
 DB 571 YLRMDDEPIAAVAFACGLATLEFVFIYRPTPVKSSSRRLCYIILAGIGLYIC 630
 OY 618 VPVYVGPVSTCLOALFPLCTICISCAVRSFQVCAF-----KMASFRPAYSVM 672
 DB 631 TFCILAKKQIYCYIQRIGIGISPMASYSALVTKNRIARILAGSKKICRKPFRMSAC 690
 OY 673 VRYOGPVYMAFIV--LKMIVIVIGMLARPOSHPTDDPKRTITVSCNPNYRN--SLFN 730
 DB 691 AQ-----LVIAFLICIGIIGIILAFIMEP-----PD-----IMHYPSIREVYLICN 733
 OY 731 TSDLLLSVGF-----FAYMKEPLPNINEAKFITLSMTFYTSVSLCTEYSA 781
 DB 734 TTNLGVVPLPGNGLIILSCFVAFKTRNVANPANEAKYIFTM--YTTCIMLAFVPI 790
 OY 782 YSGVLVTVDLTVLNLALISIG--YEGPKCYMLFPERNPAFENS 828
 DB 791 YFGSNYKILIMCFVSLSATVAGCMFVPKYIILAKPERNRSAPFT 838

RESULT 15
 MGRS_HUMAN STANDARD; PRT; 1212 AA.
 AC P41594;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metabotropic glutamate receptor 5 precursor (mglur5).
 GN GRM5 OR GPRC1E OR MGLUR5.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID:9606;
 [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=94197696; PubMed=7908515;
 RA Minakami R., Katsuki F., Yamamoto T., Nakamura K., Sugiyama H.;
 RT "Molecular cloning and the functional expression of two isoforms of
 human metabotropic glutamate receptor subtype 5";
 RL Biochem. Biophys. Res. Commun. 199;1136-1143(1994).
 [2]
 RP REVISIONS.
 RA Katsuki F.;
 RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.
 [3]
 RP SEQUENCE OF 860-952 FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=93343913; PubMed=7688218;
 RA Minakami R., Katsuki F., Sugiyama H.;
 RT "A variant of metabotropic glutamate receptor subtype 5: an
 evolutionally conserved insertion with no termination codon";
 RL Biochem. Biophys. Res. Commun. 194;622-627(1993).
 [1]
 RP FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS
 MEDIATED BY A G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-
 CALCIUM SECOND MESSENGER SYSTEM AND GENERATES A CALCIUM-ACTIVATED
 CHLORIDE CURRENT.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: 5A (SHOWN HERE) AND 5B: ARE
 PRODUCED BY ALTERNATIVE SPLICING. THEY DIFFER BY AN INSERTION OF
 32 RESIDUES.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST: TO MGLUR1.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; D28538; BA05891.1; -;
 DR EMBL; D28539; BA05892.1; -;
 DR EMBL; S64316; AAD13954.1; -;
 DR Genew: HGNC:4597; GRM5.
 DR MIM; 604102; -;
 DR InterPro; IPR001828; ANF_receptor.
 DR InterPro; IPR000337; GPCR_Mgr.
 DR Pfam; PF00003; 7tm_3; 1.
 DR Pfam; PF01094; ANF_receptor; 1.
 DR PRINTS; PR00248; GPCRMR.
 DR PROSITE; PS00979; G_PROTEIN_RECP_F3_1; 1.
 DR PROSITE; PS00980; G_PROTEIN_RECP_F3_2; 1.
 DR PROSITE; PS00981; G_PROTEIN_RECP_F3_3; 1.
 DR PROSITE; PS0259; G_PROTEIN_RECP_F3_4; 1.
 KW G-protein coupled receptor; Transmembrane; glycoprotein; signal;
 KW Multigene family; Alternative splicing.
 FT SIGNAL 1 20
 FT CHAIN 21 1212 METABOTROPIC GLUTAMATE RECEPTOR 5.
 FT DOMAIN 22 579 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 580 602 I (POTENTIAL).
 FT DOMAIN 603 616 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 617 637 II (POTENTIAL).
 FT DOMAIN 638 648 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 649 667 III (POTENTIAL).
 FT DOMAIN 668 693 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 694 714 IV (POTENTIAL).
 FT DOMAIN 715 737 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 738 759 V (POTENTIAL).
 FT DOMAIN 760 772 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 773 795 VI (POTENTIAL).
 FT DOMAIN 796 801 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 802 827 VII (POTENTIAL).
 FT DOMAIN 828 1212 VIII (POTENTIAL).
 FT CARBOHYD 88 88 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 210 210 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 378 378 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 734 734 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT VARSPLIC 877 908 MISSING (IN ISOFORM 5A).
 SQ SEQUENCE 1212 AA; 132468 MW; A3C7360668106A25 CRC64;

Query Match 15.2%; Score 675.5; DB 1; Length 1212;
 Best local Similarity 25.7%; Pred. No. 3.9e-42;
 Matches 230; Conservative 153; Mismatches 333; Indels 179; Gaps 34;

QY 28 YLPDYLGLGLFSLHANMKGIVHLNLEQVPMCKEYEVK-----IGYNLQAMREFAVEE 81
 DB 30 HMPGDIIIGALFVSF-----H-----QPTVDYHERKCGAVREQYGIQREVEAMLHTLER 78
 QY 82 INDDSLIPGVLLGYELVDYCYISN-NVOPVLYFLAHEDNLPLOED-----YSN 130
 DB 79 INSDPTLLPITIGCEIRDCMWSAVALLEQSIIEFI--RDSLISEEELGVRCDGSSSS 136
 QY 131 YISR--VAVIGPDNSESVTVANFLSLFLLPOTYSASIDELDKVFPALLTTPSPAD 188
 DB 137 FRSKKPIYVIGIGSSSVALQVONLQLENIPIQAIYATSMDSKTLFFKYPVPSDA 196
 QY 189 HVEAMVQLMHRMMNIIVLVSSDYGRDNGQLGERVARDCIAFOETLPTLPON 248
 DB 197 QQARAMVDIVKRYNMTVVSAYHTEGNEGSEAFKMSAKEGICIAHSYKIYSNAGEOS 256
 QY 249 MTSEERRLTYIDKLOOS-----TARYVVSPTLTILHFENEVLRONTGA-VVIAS 301
 DB 257 F-----DKLKLKLTSHLPKARYVACFCGEMTVRGLMLAMRRLGLAGEFILLGS 304
 QY 302 ESWAIDPVNLHNLTELGHTLGTIT--QSVPI-----PGSEFEREW 341
 DB 305 DGMW--DRYDVTGQREAVGGITIKQSPDYKMFDDYLYKLREPENHNHPMFEFMOH 361
 QY 342 GPQAGPPPLSRTSQSTYCNQECNCLNATLSFNTILRSGERYYS---VYSAVYAVAH 397
 DB 362 RFQCRLEGFOENSKY--NKTTC-----NSILKLTGHVODSKMGFVINALIYSMAY 409
 QY 398 ALHSILGCKSTC-----TRVYVPMQLLEIMKVNFT-LLDHOFEDPOGDVALHL 448
 DB 410 GLHNW--QMSLPGYAGLGDAMKPIDGRKLSLMKTNFTVSGDTILFDENGDSRGY 466
 QY 449 EIVOMQ-----WDRSONFQSVASYYPLORLKNIDISMTVNTIPMSM 494
 DB 467 EIMNFKEMGKDYFDYINVSMDNG-----ELKMDDEYWSKSNIT--RSV 510
 QY 495 CSKRCSGQ---KKRPVGIHVCCFECIDCLPGTFLNHTEDXEYEQACPNNEWSYOSETS 550
 DB 511 CSBPCEKGQIKVIRKGEVS---CCWCTCPCKENEYV---FDEYTCRACQQLGSMWPTDULTG 564
 QY 551 CFRQLVTEFMHEAPTIJAVALLAAGLSTLALIVFMRRPOTPIYRSAGBPCFLMLTL 610
 DB 565 CDLIPQVYLKMGDEPPIAAVAVFACGLGLATLEVTFVFIIRDPVYKSSSRRELCYIILAG 624
 QY 611 LTVAVVVPVYVGPVKVSTCLCRQALEPFLCTTICISCIAYRSFOIYCAF-----KMASRF 665
 DB 625 ICIQYICTPCLIAKPKQIYCYLQIRIGIGLSPMSYSALVTKTNRIARIILAGSKKICTK 684

QY 666 PRASYVWRYGYPVSNMFTTV-LKMYIVVIGMLARPOSHRPDDPKITIVSCNPNYR 724
 DB 685 PRFMSACQ-----LVIAFLICIQIGIYALIMPE-----PD-----IMHDPISR 727
 QY 725 N-SILENTSLDLLISVGS-----FAYNGKELPTVYNBAKFTTSLMTYFTSSVS 774
 DB 728 EYVLICWTNMGVVTPLGYNGLLILSCTFYAFKTRNVNPNANEAKYIAFTM---YTTCL 784
 QY 775 LCTFMSAGVLTIVDLYVNLNLAISLG-YEGPCYMLFPERENTPAYFNS 828
 DB 785 WIAFVPIYFSNKKIITMCFSVLSLSTVALGCMFVRKYVITILAKPERNVSATFT 839

Search completed: May 19, 2003, 09:50:27
 Job time : 15.9089 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:48:15 ; Search time 38.1811 Seconds
(without alignments)
4522.333 Million cell updates/sec

Title: US-09-927-315-9
Perfect score: 4443
Sequence: 1 MGPRAKTICSEFLMWLA.....ERNPAYENMIGCYTMRD 838

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs., 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

SPTREMBL_21:*
1: sp_archaea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_virus:*
16: sp_bacteriap:*
17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4392.5	98.9	839	4 O8TE23	O8TE23 homo sapien
2	3203.5	72.1	843	11 O9Z0R7	O9Z0R7 rattus norv
3	3163.5	71.2	843	11 O9Z514	O9Z514 mus musculu
4	3151.5	70.9	843	11 O9Z3J8	O9Z3J8 mus musculu
5	1454	32.7	840	11 O9Z0R8	O9Z0R8 rattus norv
6	1441	32.4	842	11 O9Z515	O9Z515 mus musculu
7	1440	32.4	842	11 O99PG5	O99PG5 mus musculu
8	1434	32.3	842	11 O9Z3J9	O9Z3J9 mus musculu
9	1431	32.2	842	11 O99PG6	O99PG6 mus musculu
10	1253.5	28.2	763	4 O8TDJ9	O8TDJ9 homo sapien
11	1062.5	23.9	940	13 O73635	O73635 fugu rubrip
12	1039	23.4	858	11 O9Z3K1	O9Z3K1 rattus norv
13	1030	23.2	940	13 O90WL6	O90WL6 spatus aura
14	1018	22.9	858	11 O9Z5D8	O9Z5D8 mus musculu
15	1014	22.8	858	11 O9Z5A4	O9Z5A4 mus musculu
16	1013	22.8	858	11 O9Z3K0	O9Z3K0 mus musculu

17	1012.5	22.8	877	13 O9PW88	O9PW88 carassius a
18	1012	22.8	858	11 O9Z5D9	O9Z5D9 mus musculu
19	1012	22.8	858	11 O91VH4	O91VH4 mus musculu
20	980	22.1	864	13 O73637	O73637 fugu rubrip
21	975.5	22.0	880	13 O73639	O73639 fugu rubrip
22	925	20.8	875	13 O73640	O73640 fugu rubrip
23	911	20.5	868	13 O73636	O73636 fugu rubrip
24	907.5	20.4	848	13 O93553	O93553 carassius a
25	898	20.2	856	13 O73638	O73638 fugu rubrip
26	824	18.5	912	11 O70410	O70410 mus musculu
27	822	18.5	844	13 O93552	O93552 carassius a
28	744.5	16.8	779	11 O35269	O35269 rattus norv
29	718	16.2	879	11 O90XS2	O90XS2 mus musculu
30	710	16.0	877	4 O8TBH9	O8TBH9 homo sapien
31	693.5	15.6	850	11 O35189	O35189 mus musculu
32	693.5	15.6	983	11 O6Z916	O6Z916 rattus norv
33	671.5	15.1	1156	13 O98UC6	O98UC6 gallus gall
34	671.5	15.1	1188	13 O98UC5	O98UC5 gallus gall
35	671.5	15.1	1242	13 O98UC4	O98UC4 gallus gall
36	665.5	15.0	855	11 O70409	O70409 mus musculu
37	654	14.7	1199	11 O9EPV6	O9EPV6 mus musculu
38	653	14.7	866	11 O35268	O35268 rattus norv
39	649	14.6	1218	13 O902F3	O902F3 oncorhynch
40	644.5	14.5	803	11 O35191	O35191 mus musculu
41	636.5	14.3	977	13 O9PWE1	O9PWE1 ictalurus p
42	632.5	14.2	1267	5 O93564	O93564 caenorhabdi
43	617.5	13.9	457	4 O8TDH8	O8TDH8 homo sapien
44	607.5	13.7	808	11 O35190	O35190 mus musculu
45	566	12.7	852	11 O35192	O35192 mus musculu

ALIGNMENTS

RESULT 1

ID O8TE23 PRELIMINARY; PRF; 839 AA.
AC O8TE23:
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Taste-specific G-protein coupled receptor TIR2.
GN TASTIR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21927605; PubMed=11917125;
RA Li X., Staszewski L., Xu H., Durick K., Zoller M., Adler E.;
RT "Human Receptors for sweet and umami taste."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4692-4696(2002).
DR EMBL; AF458149; AAM12239.1; JOINED.
DR EMBL; AF458150; AAM12239.1; JOINED.
DR EMBL; AF458151; AAM12239.1; JOINED.
DR EMBL; AF458152; AAM12239.1; JOINED.
DR EMBL; AF458153; AAM12239.1; JOINED.
KW Receptor.
SQ SEQUENCE 839 AA; 95072 MW; 8CED6F1503F05FDE CRC64;

Query Match 98.9%; Score 4392.5; DB 4; Length 839;
Best Local Similarity 99.3%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
OY 1 MGPRAKTICSEFLMWLAEPAPNSDFYLPGLGGLFSLHANKGYHLNLFLOVPMCK 60
DB 1 MGPRAKTICSEFLMWLAEPAPNSDFYLPGLGGLFSLHANKGYHLNLFLOVPMCK 60
OY 61 EYEYVIGYNLMQAMRFAVEETINDSSLLPGVLLGAEIVDVCYISNNVQVLYFLAHEDN 120
DB 61 EYEYVIGYNLMQAMRFAVEETINDSSLLPGVLLGAEIVDVCYISNNVQVLYFLAHEDN 120

QY 121 LLEIODEYSNYSRVAVIGPDNSESVMYANFLSLFLPQITYSAISDELDRKVRPAL 180
 DB 121 LLEIODEYSNYSRVAVIGPDNSESVMYANFLSLFLPQITYSAISDELDRKVRPAL 180
 QY 181 LRTTSPADHVEAMVQMLHFRNMWIIYVSSDTYGRDNGOLGERRARDICIAFOETL 240
 DB 181 LRTTSPADHVEAMVQMLHFRNMWIIYVSSDTYGRDNGOLGERRARDICIAFOETL 240
 QY 241 PTLQPNQNMNTSEERQRLVTVYDKLOOSTARVAVYVSPDLTYHFENYLRONFGAWIA 300
 DB 241 PTLQPNQNMNTSEERQRLVTVYDKLOOSTARVAVYVSPDLTYHFENYLRONFGAWIA 300
 QY 301 SESMAIDPVNLHNLTELGLTGLTITQSVPIPGSEFEREMQAGPPPLSRTSQTG 360
 DB 301 SESMAIDPVNLHNLTELGLTGLTITQSVPIPGSEFEREMQAGPPPLSRTSQTG 360
 QY 361 QECNCLNATLSFNTILRLSGERVVSVYSAVVAHALHSLGCDSTCKRKYVPMQ 420
 DB 361 QECNCLNATLSFNTILRLSGERVVSVYSAVVAHALHSLGCDSTCKRKYVPMQ 420
 QY 421 LLEIMKVNFTLLDHOIFEDPOGVALHLEIYOMQDRSONFQSVASYPLQROLKNTOD 480
 DB 421 LLEIMKVNFTLLDHOIFEDPOGVALHLEIYOMQDRSONFQSVASYPLQROLKNTOD 480
 QY 481 ISWHTVNTIPMSKSRCSGQKKRPVGIHVCCECICDLPCTFLNTEDEYECQACP 540
 DB 481 ISWHTVNTIPMSKSRCSGQKKRPVGIHVCCECICDLPCTFLNTEDEYECQACP 540
 QY 541 NEMYSQSETSCFRQLVLEHHEAPTAVALALGFLSTALIVYMRHQTIVSAG 600
 DB 541 NEMYSQSETSCFRQLVLEHHEAPTAVALALGFLSTALIVYMRHQTIVSAG 600
 QY 601 GPMCFMLTLLVAVYVYVGPVKYSTCICRQALPPLCTICISCIARVFOYCAFK 660
 DB 601 GPMCFMLTLLVAVYVYVGPVKYSTCICRQALPPLCTICISCIARVFOYCAFK 660
 QY 661 MASFPRAYSYVWRYOGPYVSMATYVVKWYIYIGMLARPOS--HPTDPDPKITYSC 719
 DB 661 MASFPRAYSYVWRYOGPYVSMATYVVKWYIYIGMLARPOS--HPTDPDPKITYSC 719
 QY 720 NPNRNSLLENTSLDLLSVGFSFAVAGKELPTNYNEAKITLTMPEYFSSVSLCFM 779
 DB 720 NPNRNSLLENTSLDLLSVGFSFAVAGKELPTNYNEAKITLTMPEYFSSVSLCFM 779
 QY 781 SAYSGVLVTVYDLVTVLNLALISLGYFGPKCYMILFPERNTPAYFNSMIOGYTMRD 839
 DB 781 SAYSGVLVTVYDLVTVLNLALISLGYFGPKCYMILFPERNTPAYFNSMIOGYTMRD 839

RESULT 2
 ID 0920R7 PRELIMINARY; PRT; 843 AA.
 AC 0920R7;
 DT 01-MAY-1999 (Tremblrel. 10, Created)
 DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
 DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
 DE Putative taste receptor tr2 (Fragment).
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniala; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
 OC NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=K1STAR;
 RX MEDLINE=99159821; Pubmed=10052456;
 RA Hoon M.A., Adler E., Lindemeier J., Battey J.F., Ryba N.J.,
 RA Zuker C.S.;
 RT "Putative mammalian taste receptors: a class of taste-specific GPCRs
 RT with distinct topographic selectivity.";
 RL Cell 96:541-551(1999).
 DR EMBL: AF127390; AAD18070.1; -
 DR InterPro: IPR001828; ANF_receptor.

DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00259; G_PROTEIN_RECP_F3_4; 2.
 KW Receptor.
 FT NON_TER
 SQ SEQUENCE 843 AA; 95799 MW; D23AC22D1E049B8 CRC64;
 Query Match 72.1%; Score 3203.5; DB 11; Length 843;
 Best Local Similarity 70.6%; Pred. No. 1.3e-257;
 Matches 395; Conservative 108; Mismatches 133; Indels 7; Gaps 4;
 QY 1 MGPRAKTICSLFLFLWYLAEP--AENSDFYLPDGYLLGLFSLHANKGIVHLNFKQVP 57
 DB 1 MGPRAKTICSLFLFLWYLAEP--AENSDFYLPDGYLLGLFSLHANKGIVHLNFKQVP 57
 QY 58 MCKREYKVGYNIMQAMRAVEEINDSSILPEVLLGLEYVDVGYCISNNVQPLVYLAH 117
 DB 61 KNEFTKRVLGYNIMQAMRAVEEINNCSLLPEVLLGLEYVDVGYCISNNVQPLVYLAH 120
 QY 118 EDNLPLEIODEYSNYSRVAVIGPDNSESVMYANFLSLFLPQITYSAISDELDRKVRP 177
 DB 121 DDLLPLTKDYSQYMPRAVAVIGPDNSESATYNSIISHLIPQITYSAISDELDRKVRP 180
 QY 178 PALLRTTSPADHVEAMVQMLHFRNMWIIYVSSDTYGRDNGOLGERRAR-RDICIAP 236
 DB 181 PSMRLTPVSAHHEAMVQMLHFRNMWIIYVSSDTYGRDNGOLGERRAR-RDICIAP 240
 QY 237 QETLPTQPNQNMNTSEERQRLVTVYDKLOOSTARVAVYVSPDLTYHFENYLRONTGA 296
 DB 241 QETLPTQPNQNMNTSEERQRLVTVYDKLOOSTARVAVYVSPDLTYHFENYLRONTGA 300
 QY 297 VWIASEMAIDPVNLHNLTELGLTGLTITQSVPIPGSEFEREMQAGPPPLSRTSQTG 356
 DB 301 VWIASEMAIDPVNLHNLTELGLTGLTITQSVPIPGSEFEREMQAGPPPLSRTSQTG 360
 QY 357 YTCNQECNCLNATLSFNTILRLSGERVVSVYSAVVAHALHSLGCDSTCKRKYV 416
 DB 361 YTCNQECNCLNATLSFNTILRLSGERVVSVYSAVVAHALHSLGCDSTCKRKYV 420
 QY 417 PMQLLEIMKVNFTLLDHOIFEDPOGVALHLEIYOMQDRSONFQSVASYPLQROLK 476
 DB 421 PMQLLEIMKVNFTLLDHOIFEDPOGVALHLEIYOMQDRSONFQSVASYPLQROLK 480
 QY 477 NIODISWHTVNTIPMSKSRCSGQKKRPVGIHVCCECICDLPCTFLNTEDEYECQ 536
 DB 481 YINNVSMYTPRNTYVPMSCSKSCQPGQMKRSVGLHPCCFECLDQMPGTIYLRNSADENCL 540
 QY 537 ACNNENMSYQSETSCFRQLVLEHHEAPTAVALALGFLSTALIVYMRHQTIVSAG 596
 DB 541 SCPSGMSYKNDITCFORRPLEFLEHHEAPTAVALALGFLSTALIVYMRHQTIVSAG 600
 QY 597 RSAGGPMCFMLTLLVAVYVYVGPVKYSTCICRQALPPLCTICISCIARVFOYCAFK 656
 DB 601 RSAGGPMCFMLTLLVAVYVYVGPVKYSTCICRQALPPLCTICISCIARVFOYCAFK 660
 QY 657 CAFEMASRFPRAYSYVWRYOGPYVSMATYVVKWYIYIGMLARPOS--RTDPDPKIX 714
 DB 661 CAFEMASRFPRAYSYVWRYOGPYVSMATYVVKWYIYIGMLARPOS--RTDPDPKIX 719
 QY 715 TVYSCNPNRNSLLENTSLDLLSVGFSFAVAGKELPTNYNEAKITLTMPEYFSSV 774
 DB 720 TVYSCNPNRNSLLENTSLDLLSVGFSFAVAGKELPTNYNEAKITLTMPEYFSSV 779
 QY 775 LCTFMSAYSGVLVTVYDLVTVLNLALISLGYFGPKCYMILFPERNTPAYFNSMIOGYT 834
 DB 780 LCTFMSAYSGVLVTVYDLVTVLNLALISLGYFGPKCYMILFPERNTPAYFNSMIOGYT 839
 QY 835 MRR 837
 DB 840 MKR 842

RESULT 3	ID	Q92514	Q92514	PRELIMINARY:	PRT:	843 AA.
AC	Q92514					
DC	Q92514					
DT	01-DEC-2001	(TREMBLrel. 19, Created)				
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)				
DT	01-JUN-2002	(TREMBLrel. 21, Last annotation update)				
DE		Candidate taste receptor T1R2.				
GN		TAS1R2 OR T1R2.				
OS		Mus musculus (Mouse).				
OC		Eukaryota; Metazoa; Chordata; Cranialta; Vertebrata; Euteleostomi;				
CC		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
OX		NCBI_TaxID=10090;				
RP		[1]				
RN		SEQUENCE FROM N.A.				
RC		STRAIN=CS7BL/6J;				
RX		MEDLINE=21219400; PubMed=11319557;				
RA		Montmayeur J.P., Liberles S.D., Matsunami H., Buck L.B.;				
RT		"A candidate taste receptor gene near a sweet taste locus.";				
RL		Nat. Neurosci. 4:492-498(2001).				
DR		EMBL; AF337041; AA34938.1; ..				
DR		MGI; MGI:1933546; Tas1r2.				
DR		InterPro; IPR001828; ANF_receptor.				
DR		InterPro; IPR000337; GPCR_MgF.				
DR		Pfam; PF00003; 7tm.3; 1.				
DR		Pfam; PF01094; ANF_receptor. 1.				
DR		PROSITE; PS50259; G_PROTEIN_RECPT_P3_4; 1.				
KM		Receptor.				
QO		SEQUENCE. 843 AA; 95735 MW; 0543834EA4E7AC2E CRC64;				

Query Match	71.2%	Score 3163.5	Length 843
Best Local Similarity	69.2%	Pred. No. 2.8e-254	
Matches	583	Conservative 141	Indels 7
		Mismatches 112	Gaps 4

[illegible]

Db	541	SCPGSMYSKNNINACFKRRLAFLEMHHEVPIVYITLALGAFISPLATILLFMHHPQPMV	600
Qy	557	RSAGGPMCFMLTLTLVLAAYMYPPYVGPBKVSTLCQALFPLCFTTCICICIAVRSQIY	636
Db	601	RSAGGPMCFMLVPLLELAFGVAAPYVBPPTFSCGCRQAEFTYCSVLCSITVRSQIY	660
Qy	657	CAFKKASHRPRAYSYVWVYQGPYVSMAFIVLKKVIVLVIQMLARQSHP--RTDDDPKI	714
Db	661	CYFKMARRLPSAYGFMWRHGGPVYFAETPAVKKALVAGNMKL--TTINPIGRDDPPNI	719
Qy	715	TIYSCNPNYKRNLSLFPNLSLDLLSVYGFSRAYMCKELPTYNNAKFTTILSMTEYFSSVS	774
Db	720	TIISCHPNRYKGLLEFNFSMDLLSLVGEFSRAYMCKELPTYNNAKFTTILSMTEFSSIS	779
Qy	775	LCFEMSAYSGLVITIVDLLVTLVLMILASIGYFGPKCYMLLEFPERNTPAYFNSMIQGYT	834
Db	780	LCFEMSAYHDGLVITIMDLLVTLVNLFIAGIGYFEPKCYMLLEFPERNTSAYFNSMIQGYT	839
Qy	835	MRR 837	
Db	840	MK 842	

RESULT 4

ID Q923J8 PRELIMINARY; PRT; 843 AA.

DT 01-DEC-2001 (Tremblay, 19, Created)

DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Sweet taste receptor M1R3

OS Mus musculus (Mouse).

0C Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
0X NCBI TaxID=10090:

RP SEQUENCE FROM N.A.

RA NELSON C.S.; HOOB M.A.; CHANDASHEKAR C.; ZHANG L.; AYDA N.O.F.
RA ZUKER C.S.;

Cell 0:0-0(2001).

DR InterPro; IPR001828; ANF_receptor.

DR Pfam; PF00003; 7tm_3; 1.

DR PROSITE; PS50259; G_PROTEIN_RECEP_F3_4; 1.

SQ SEQUENCE. 843 AA; 95752 MW; FD3C6B9B23270DA3 CRC64;

Query Match	70.9%	Score 3151.5	DB 11	Length 843
Best Local Similarity	68.9%	Pred. No. 2.8e+253		
Matches 581	Conservative 113	Mismatches 142	Indels 7	Gaps 4

Qy	1	MGPAKFTCSLFFLLMWLAEP----	AENSDFLPEDYDILGGFSJHAKMKIVHLENFQVP	57
Qy	1	MGPAKFTCSLFFLLMWLAEP----	AENSDFLPEDYDILGGFSJHAKMKIVHLENFQVP	57
Db	1	MGPAKFTCSLFFLLMWLAEP----	AENSDFLPEDYDILGGFSJHAKMKIVHLENFQVP	57
Db	1	MGPAKFTCSLFFLLMWLAEP----	AENSDFLPEDYDILGGFSJHAKMKIVHLENFQVP	57
Qy	58	MCKEYEVKIVGYNLMQAMRFAVEEIN	NDSSLLPGVLLGYELVDVCYSNNVQVLYFLAH	117
Db	61	KCNENKMKVLGYNLMQAMRFAVEEIN	NDSSLLPGVLLGYELVDVCYSNNVQVLYFLAH	120
Qy	118	EDNLLPFOEDXSNTISRVAANYIGD	NSWYTVANFLSLLPQITYSAISDELPOKVR	177
Db	121	IDDFPLPIKDXSQYRPOVVAAYIGD	NSWYTVANFLSLLPQITYSAISDELPOKVR	180
Qy	178	PALLRTPSADHNEAMVQLLHFRNMW	IIVLVSSDTPYGRDNGOLIGERYARR-DICIAF	236
Db	181	PALLRTPSADHNEAMVQLLHFRNMW	IIVLVSSDTPYGRDNGOLIGERYARR-DICIAF	240
Qy	237	QETLPLPQNONMTSEEROLVTVDKL	KLOOSTARVYVSSPDLTLYHFEVEYLRONFCA	296
Db	241	QETLPLPQNONMTSEEROLVTVDKL	KLOOSTARVYVSSPDLTLYHFEVEYLRONFCA	300

QY	297	WIASSEAIIDPVLTNLTELGLHCPPIGTTQSPVDPGESESEBMRGQACGPLPRLTSQS	3356
Db	301	WIASSEAIIDPVLTNLTELHNTGTFLVATIQKVSIFGESEFVRHDKPGTMAPNEISLR	3606
QY	357	YTCNOCDCNCLNATLSFNTILRLSGERVVSVSAYVAVAHAHSLGLCDKSTCTKRVY	416
Db	361	YTCNOCDCACNMTESFNNVLTMLSGERVVSVYSAYVAVAHTLRLLHCNOVRCRQIYV	4206
QY	417	PMQLLEEIMKYNFTLLDHIQIFPDQGVALHLEIVQNMQRSPQPSVASYPDLQRLK	476
Db	421	PMQLLEIREIMHNFETLLGQIQLFEDQGDPMPLDIIOMQGLSQNPFGSIASYPETRLT	4806
QY	477	NIODISMTVTNNTIPMSKSRCSGQKKKPVGIHCCFECIDCLPCTFIMHTEDEYEQ	5366
Db	481	YISNVSMTPTNNTVVISKSCSCQPGQKKRIGIHPCCFECVDCPPTIYNRSVDENCL	5406
QY	537	ACPNNMESVYQSESTCEKRLQVLEMHNEAPTIAVALLAELSTALIVTFWRHPQPIV	5966
Db	541	SCPQSMGSMYKNNINACFKRRLAFLEWHVEVPTIVTIALGPISTLAILLFMRHFGPBMV	6006
QY	597	RSAGGPKCFMLTLLIAYMYVPVYVGPVKSTLCRLQALFPLCTFICISIAVRSQIV	6566
Db	601	RSAGGPKCFMLVPLLLAFMGVYVGPPTVFCSCFRAPETVCFSCVLSCTIVRSFOIV	6606
QY	657	CAFKKASFEPRVSVFWVYQGPYSMAETVTKLVIVIGLARPQSHP--RTDPDDPKI	7146
Db	661	CVFKKARLLPSAIGFWMRYHGPYFVAFITVAKVALAAGNMILA--TTINPIGRTDPPDNI	7196
QY	715	TVSCNPYRNSLLENTSLDILLVSYGFSFAYMKELEPTYNENAKFTLLSMTEFYTSVS	7746
Db	720	ILTSCHPNYRNGLLENTSMDLLLSVLGFSFAYMKELEPTYNENAKFTLLSMTEFSYSSIS	7796
QY	775	LCTEMSAASGVLTNYVDLLVNLATSLGSEDPKCYMILFYDERNTPRAYFNSMIGYT	8346
Db	780	LCTEMSAASGVLTNYVDLLVNLATSLGSEDPKCYMILFYDERNTSAFNSMIGYT	8396
QY	835	MRR 837	
Db	840	MRR 842	
RESULT 5			
Q920R8			
ID	Q920R8	PRELIMINARY;	PRT; 840 AA.
AC	Q920R8		
DT	01-MAY-1999	(TREMBLrel. 10, Created)	
DT	01-MAY-1999	(TREMBLrel. 10, last sequence update)	
DT	01-JUN-2002	(TREMBLrel. 21, last annotation update)	
DE	Putative taste receptor Trl (Fragment).		
OS	Rattus norvegicus (Rat).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.		
OX	NCBI_TaxID=10116;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=MISTAR;		
RX	MEDLINE=99159821; PubMed=10052456;		
RA	Hoon M.A., Adler E., Lindemeier J., Battey J.F., Ryba N.J.,		
RA	Zuker C.S.;		
RT	"Putative mammalian taste receptors: a class of taste-specific GPCRs		
RT	with distinct topographic selectivity.";		
RL	Cell 96:541-551(1999).		
DR	EMBL: AF127389; AADI8069.1; -		
DR	InterPro: IPR001828; ANF_Receptor.		
DR	InterPro: IPR000337; GPCR_Mgt.		
DR	Pfam: PF00003; 7tm_3; 1.		
DR	Pfam: PF01094; ANF_receptor; 1.		
DR	PRINTS: PR00248; GPCRMR.		
DR	PROSITE: PS50259; G_PROTEIN_RECIP_F3_4; 2.		
FW	Receptor.		
FT	NON_TER	840	840
SO	SEQUENCE	840 AA;	93496 MW; 1FCFB7EFC6B845DB CRC64;

[illegible]

OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 RX MEDLINE=21219400; PubMed=11319557;
 RA Montmayeur J.P., Liberles S.D., Matsunami H., Buck L.B.,
 "A candidate taste receptor gene near a sweet taste locus,"
 RL Nat. Neurosci. 4:492-498(2001).
 DR EMBL; AF337040; AK39437.1; -.
 DR MGI; MGI:1927505; Tas1r1.
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000345; CyTC_heme_bind.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS50259; G_PROTEIN_RECER_F3_4; 1.
 KW Receptor.
 SQ SEQUENCE 842 AA; 93471 MW; FDBA0CC1BD45DEB7 CRC64;

Query Match 32.4%; Score 1441; DB 11; Length 842;
 Best Local Similarity 39.7%; Pred. No. 5.4e-111;
 Matches 328; Conservative 123; Mismatches 339; Indels 36; Gaps 15;

OY 27 FYLPDGYLLGGLFSLHANKGIVHNLFOVPMC-KEYEYKVIGYINLQMAFRVVEINND 85
 DB 33 FSLPDPFLLAGLSLHADLOVRHRL--VTSCDSDSGNGHGYLFQAMRTVEINNS 90
 OY 86 SLLPGVLLGEIVDYCYISNNVOPVLYELAHE-DNLPIQEDYSNYISRYVAVIGPDNS 144
 DB 91 TALLPNTLTGYELDYVCSSESSNVYATRLVLAQOGGHLEMDLNRHSSKVALLGPDNT 150
 OY 145 ESMVTVANFLSLFLPQITTSASIDELDKVRFALLRTTPSADHVEAMVOLMHLFRNN 204
 DB 151 DHAVTTALLSPFLMPLVSYEASVILSGKRKFPFLRTIPSDKQOVEIVYRLQSFQV 210
 OY 205 WTIIVSSDTYGRDNGQLGERVARRDICIAFOETPLTQPNQNTSEEROLVITVOKL 264
 DB 211 WISLVGSDGYQLGVQALAEELATPRGICVAFKDVPL---SAQAGDPRMOMML---RL 264
 OY 265 QOSTARVVVFPDITLYHFNENEVLRONFTGAVWIASSEMAIDPVLHNLTEGLHGTFLG 324
 DB 265 AARATTVVVFENRHLAGVFRSVLANLTGKWLASEDMAISTYITVNPQIGIGTVLG 324
 OY 325 ITIQSVPIPGSEFEREMPOA--GPPPLSRISQSTYCNQECNCLNATLSFTITRLSGE 382
 DB 325 VAIQORQVPGLEKEFEESYVQAVMGPRTCPGSGMGTNOLCRECHAFITWNNPDELGAFSM 384
 OY 383 RVVSVGVAVVVAHVALHSILGCDKSTCTKRVVYPMOLLEITWKNFTLLDQIFPDPOG 442
 DB 385 SAAYVAVVEAVVAVAGHQLGCTGCTGARGVYPMOLLOQIKYKNFLHKTVAFPDQK 444
 OY 443 DYALHLEIVQWQDRSQNPQSV--ASYPIQOROLKNIQDISWHTVNTTIPMSKSRQ 500
 DB 445 DPLGYDDIIMAMNGPEWTFEIVGASLSPLHDI-NKTKIQMHGKNQVPSVCTRDL 503
 OY 501 SQGKKRPVGHVCECDICLPGETLNFHDEDECOACPNMWSQSELSCKRQVLE 560
 DB 504 BHHNLVMSHRCFECMCEAGTFLN-TSELHTQPCGTEWMAEGSSACSRVVEFG 562
 OY 561 WHEAPTLIVALLAALGFTSTALIVF-----NRHQPPIVRSAGGPMCLMTLLILY 613
 DB 563 WHE--PISIVLLAA-----NTLLILLIIGAGLFAMR-LHTPVRSAGGLCTLMGLSLVA 615
 OY 614 AYVVVPVYVGPVSTCLQALPPLCTICISIAVRSFOIVCAFKNASRPRAVSYV 673
 DB 616 GSCSYSEFGKPTVPACILROPLFSLGFAIFLSCLTIRSFQVILFKFSTKVPPTYHTWA 675
 OY 674 RQGGVYSMAFTIVLKNYIVVIGMLARQSHRTDPD---DKITIVSCNPNYRNSLFLN 730
 DB 676 QNHGAGI---FVIVSSVHLLFCLTWLAMPRTPTREYQRFPHVLLLECTEVNSVGFVA 732

OY 731 TSLDLLSVGSPFAYMGEKLPFNNEAKFITLSMTEFTSSVSLCTEMSAYSVLYTIV 790
 DB 733 FAHNILLSTVFCVSYLGELEPENTNEAKVCFSLHLPVSVIATFTMSITQGSILP 792
 OY 791 DLLVTVNLNLAISLGFGRKCYMILFPERNPAPVAFNSMIGCYTKR 836
 DB 793 NVLAGLATLNSGSGFLPKCYVILCPRLNTEHFPQASIQDYTRR 838

RESULT 7

O99PG5 PRELIMINARY; PRT; 842 AA.

ID O99PG5
 AC O99PG5;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
 DE Putative sweet taste receptor TIR1 (Fragment).
 GN TASIR1 OR GPR70.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RC SEQUENCE FROM N.A.
 RC STRAIN=129P3/J;
 RX MEDLINE=21030739; PubMed=11178737;
 RA Li X., Inoue M., Reed D.R., Hugue T., Puchalski R.B., Tordoff M.G.,
 Niomiya Y., Beauchamp G.K., Bachmanov A.A.;
 "High-resolution genetic mapping of the saccharin preference locus
 (Sec) and the putative sweet taste receptor (TIR1) gene (Gpr70) to
 RT mouse distal chromosome 4";
 RL Mamm. Genome 12:13-16(2001).
 DR EMBL; AF301162; AK07092.1; -.
 DR MGI; MGI:1927505; Tas1r1.
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000345; CyTC_heme_bind.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS50259; G_PROTEIN_RECER_F3_4; 1.
 KW Receptor.
 FT NON-TER 842
 SQ SEQUENCE 842 AA; 93448 MW; 39739A2FF482D33F CRC64;

Query Match 32.4%; Score 1440; DB 11; Length 842;
 Best Local Similarity 39.7%; Pred. No. 6.6e-111;
 Matches 328; Conservative 122; Mismatches 340; Indels 36; Gaps 15;

OY 27 FYLPDGYLLGGLFSLHANKGIVHNLFOVPMC-KEYEYKVIGYINLQMAFRVVEINND 85
 DB 33 FSLPDPFLLAGLSLHADLOVRHRL--VTSCDSDSGNGHGYLFQAMRTVEINNS 90
 OY 86 SLLPGVLLGEIVDYCYISNNVOPVLYELAHE-DNLPIQEDYSNYISRYVAVIGPDNS 144
 DB 91 TALLPNTLTGYELDYVCSSESSNVYATRLVLAQOGGHLEMDLNRHSSKVALLGPDNT 150
 OY 145 ESMVTVANFLSLFLPQITTSASIDELDKVRFALLRTTPSADHVEAMVOLMHLFRNN 204
 DB 151 DHAVTTALLSPFLMPLVSYEASVILSGKRKFPFLRTIPSDKQOVEIVYRLQSFQV 210
 OY 205 WTIIVSSDTYGRDNGQLGERVARRDICIAFOETPLTQPNQNTSEEROLVITVOKL 264
 DB 211 WISLVGSDGYQLGVQALAEELATPRGICVAFKDVPL---SAQAGDPRMOMML---RL 264
 OY 265 QOSTARVVVFPDITLYHFNENEVLRONFTGAVWIASSEMAIDPVLHNLTEGLHGTFLG 324
 DB 265 AARATTVVVFENRHLAGVFRSVLANLTGKWLASEDMAISTYITVNPQIGIGTVLG 324
 OY 325 ITIQSVPIPGSEFEREMPOA--GPPPLSRISQSTYCNQECNCLNATLSFTITRLSGE 382
 DB 325 VAIQORQVPGLEKEFEESYVQAVTGPRTCPGSGMGTNOLCRECHAFITWNNPDELGAFSM 384

QY 383 RYVSVYSAVYAAVAHLSLGGDSKSTCRKRVYPMOLLEIMKVNFTLLDHOIFPDG 442
 DB 385 SAAYNYEAYVAAGHGLHOLGCTSGTCARGPYPMOLLOQIKVNLHKKTYAFDNG 444
 QY 443 DVALHLEIVOMQDRSQNPQSV--ASYPLQRLKNIODISMTVNTTIPMSCKSRQ 500
 DB 445 DPLGYDDIAMDNGPMTPEVIGSASLSPVHDI--NKRITOMHGKNNQVPVSVCTRDCL 503
 QY 501 SGQKKRKGVIHVOCCFECIDLPGLFNLHTEDEYECACPNMENSYSOSETCFKRLVFE 560
 DB 504 EGHRLVMGSHCCFECMPCEAGTFILN-TSELHQCQCEEAAPBESSACFSRTVEFLG 562
 QY 561 WHEAPITVAALLAIGLSTLAIIVF-----WRHQPPIVRSAGCPMCFMLTLLV 613
 DB 563 WHE--PISVLIAA-----NTLLILLIGTAGLFAMR-LHTPVRSAGRCFLMLGSLVA 615
 QY 614 AYAVPYVYPPRYVSTCLCQALFPLCTTCISCIASVRSQIYCAEFMASRFPRAVYVW 673
 DB 616 GSCSLXSFEEPPVPAQLRQPLFSIGFAIFLSCITRSFQLYIFEFSTKVPFFYHTWA 675
 QY 674 RYGPVYMAFITVYKVIYVIGMLARQSHPRDPP--DKPTIYSCPNRNSILFN 730
 DB 676 QNHGAGI--FVIYSSIVHFLCLITWLAMTPRPTREYQRPPLVILCEEVNSVGLVA 732
 QY 731 TSLDLLSVYGFSAVYKGLPTNYNEAKFTLSMPTFTSSVSLCTFMSAGVLTIV 790
 DB 733 FAHNIILISTFVCSYLGKELPENYNAKCVTSILHFWSMIAFWMSSIIYGSYLPAV 792
 QY 791 DLVTVYNLALAIISLGYRPGCYMILFPERNTRAYFNSMIOGTMR 836
 DB 793 NVLAGLATLSGGFSGYFLPKCYVILCRPELNTNTEHFQASIODYTRR 838

RESULT 8

0923J9

ID 0923J9 PRELIMINARY; PRT; 842 AA.

AC 0923J9;
 DT 01-DEC-2001 (TReMBLrel. 19, Created)
 DT 01-DEC-2001 (TReMBLrel. 19, last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, last annotation update)
 DE Taste receptor TIR1.
 GN TASIR1.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxId=10090;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA Nelson G., Hoon M.A., Chandrasekar J., Zhang Y., Ryba N.J.P.,
 RA Zuker C.S.,
 RT "Mammalian Sweet Taste Receptors";
 RL Cell 0:0-0(2001).
 DR EMBL: AY032622; AK51603.1; -
 DR MGD; MGI:1927505; Tasir1.
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000345; CytC_heme_bind.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm.3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS50259; G_PROTEIN_RECEP_F3_4; 1.
 KW Receptor.
 SQ SEQUENCE 842 AA; 93455 MW; D4DCE90959E991A CRC64;

Query Match

Best Local Similarity 32.3%; Score 1434; DB 11; Length 842;
 Matches 327; Conservativity 39.6%; Pred. No. 2.1e-110;
 Mismatches 340; Indels 36; Gaps 15;

QY 27 FYVPGDYLGGLGSLAHNMKGIVHLNFIQVPMG-KEYEVAVIGVNLMOARPAVEEINND 85
 DB 33 FSLPGDFLGLSLAHNDCIQVRRRL--VTSCDRSDSFNGHGHIHLFQAMRFVEEINNS 90

QY 86 SSLPLGVLGEYVDYVCISNNYQVPLYFLAHE-DNLLPIQEDYSNYSIRVAVIGPDS 144
 DB 91 TALPNTLTGELYDYDCSESSNYATLRVPAQOCTGLHEMRDLRNHSSKVALLIGPDT 150
 QY 145 ESWTVANFSLLELLRQITTSASDELKDYRRPALLRTPSDHNEAVQMLHRMN 204
 DB 151 DHAVTAAALSPMLPVLVSYEASSVILSGRKRPESFLRTIPSKYQYEVIRLLQSGWV 210
 QY 205 WITVLVSSDYGNGDGLSERVARDICIAFETETLTLPONNMTEEQRLVTVYDKL 264
 DB 211 WISLVSGYGGQGVQALBELATPRKICVAFKDYVL----SAQAGPQRQRMV---RL 264
 QY 265 QOSTARVYVSPDLTYHFNELRQNTGAVWIASESAIDPVLNLTGLHGTFLG 324
 DB 265 ARARTVYVVFNSRHLAGVFRRSVLANLGKWIASEDAISYITNVGIGIGTVLG 324
 QY 325 ITIOSVPIPGFSEREPKPA--GPPLSRTSOSYTCNOECDCLNLTISFNILRSGE 382
 DB 325 VALQORQVPLKFEESYQAVMGAPRTCEGSGWCTNOLCRCHAFITVNMDELGFMSM 384
 QY 383 RYVSVYSAVYAAVAHLSLGGDSKSTCRKRVYPMOLLEIMKVNFTLLDHOIFPDG 442
 DB 385 SAAYNYEAYVAAGHGLHOLGCTSGTCARGPYPMOLLOQIKVNLHKKTYAFDNG 444
 QY 443 DVALHLEIVOMQDRSQNPQSV--ASYPLQRLKNIODISMTVNTTIPMSCKSRQ 500
 DB 445 DPLGYDDIAMDNGPMTPEVIGSASLSPVHDI--NKRITOMHGKNNQVPVSVCTRDCL 503
 QY 501 SGQKKRKGVIHVOCCFECIDLPGLFNLHTEDEYECACPNMENSYSOSETCFKRLVFE 560
 DB 504 EGHRLVMGSHCCFECMPCEAGTFILN-TSELHQCQCEEAAPBESSACFSRTVEFLG 562
 QY 561 WHEAPITVAALLAIGLSTLAIIVF-----WRHQPPIVRSAGCPMCFMLTLLV 613
 DB 563 WHE--PISVLIAA-----NTLLILLIGTAGLFAMR-LHTPVRSAGRCFLMLGSLVA 615
 QY 614 AYAVPYVYPPRYVSTCLCQALFPLCTTCISCIASVRSQIYCAEFMASRFPRAVYVW 673
 DB 616 GSCSLXSFEEPPVPAQLRQPLFSIGFAIFLSCITRSFQLYIFEFSTKVPFFYHTWA 675
 QY 674 RYGPVYMAFITVYKVIYVIGMLARQSHPRDPP--DKPTIYSCPNRNSILFN 730
 DB 676 QNHGAGI--FVIYSSIVHFLCLITWLAMTPRPTREYQRPPLVILCEEVNSVGLVA 732
 QY 731 TSLDLLSVYGFSAVYKGLPTNYNEAKFTLSMPTFTSSVSLCTFMSAGVLTIV 790
 DB 733 FAHNIILISTFVCSYLGKELPENYNAKCVTSILHFWSMIAFWMSSIIYGSYLPAV 792
 QY 791 DLVTVYNLALAIISLGYRPGCYMILFPERNTRAYFNSMIOGTMR 836
 DB 793 NVLAGLATLSGGFSGYFLPKCYVILCRPELNTNTEHFQASIODYTRR 838

RESULT 9

099PG6

ID 099PG6 PRELIMINARY; PRT; 842 AA.

AC 099PG6;
 DT 01-JUN-2001 (TReMBLrel. 17, Created)
 DT 01-JUN-2001 (TReMBLrel. 17, last sequence update)
 DT 01-JUN-2002 (TReMBLrel. 21, last annotation update)
 DE Putative sweet taste receptor TIR1 (Fragment).
 GN TASIR1 OR GPR70.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxId=10090;
 RX [1]
 RP SEQUENCE FROM N.A.
 RA STRAIN=C57BL/6BYJ;
 RC MEDLINE=21030739; PubMed=11178737;
 RA Ito X., Inoue M., Reed D.R., Huque T., Puchalski R.B., Tordoff M.G.,
 RA Nishimura Y., Beauchamp G.K., Bachmanov A.A.;
 RT "High-resolution genetic mapping of the saccharin preference locus

RT (Sac) and the putative sweet taste receptor (T1R1) gene (Gpr70) to
 RT mouse distal chromosome 4.1;
 RL Mamm. Genome 12:13-16(2001).
 DR EMBL; AF301161; AK07091.1; -
 DR MGI; MGI:1927505; Tas1r1.
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000345; Cytc_heme_bind.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam; PF01094; ANF_receptor; 1.
 DR Pfam; PF01094; ANF_receptor; 1.
 DR PRINTS; PR00248; GPCRMR.
 DR PROSITE; PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE; PS50259; G_PROTEIN_RECPT_F3_4; 1.
 KW Receptor.
 FT NON_TER
 SO SEQUENCE 842 AA; 93425 MW; 58826C43F5DD352B CRC64;

Query Match 32.2%; Score 1431; DB 11; Length 842;
 Best Local Similarity 39.6%; Pred. No. 3.7e-110;
 Matches 327; Conservative 124; Mismatches 339; Indels 36; Gaps 15;

QY 27 FYLDGDIYLGSLHANKGIVHNLQVPMG-KEYEVKIVGIMQAMRPAVEIIND 85
 DB 33 FSLPGDFLLAGLFSIHADCLQVRRHPL--VTSRSDSFNGHGYHLQAMFEVEIINS 90
 QY 86 SSLPGVLLGEIVDVCYISNNQPVLYFLAHE-DNLLPIQEDYSNYSRYVAIVGPDNS 144
 DB 91 TALPNTLTGLDYDCESSNVTATRLVLAQOGLHEMRDRDLNHSKVALLGPDNT 150
 QY 145 ESMVTANFLSLFLPQTYSAISDELNDKVRFPALLRTTSPADHIVAMVQLMHPMN 204
 DB 151 DHAVTTALLSPFLMPLEVSSEASSVILSGKRKFPFLRTISDKQVEIVIRLLQSPGV 210
 QY 205 WIIYVSSDTYGRDNGQLLGERVARRDICIAFOETLPLQPNQMTSEKRLVTIVDKL 264
 DB 211 WISLVGSYGDGLQVOALEELAPRGICVAFKNVPL--SAQAGDPRMRML--RL 264
 QY 265 QOSTRVVVVSPDLTLHFENEVLRFONFTGAVMIASEMAIDPLHLTELGHGFLG 324
 DB 265 ARARTVVVVSNNRLDGVFFRSVYLANLTGKVMIASDMAISTNIPVSGIGIYVLG 324
 QY 325 ITIOSVPIGSEFSEFREMPOA--GPPPLSRTSOSTYCNQECNCLNATLSFNTLRLSGE 382
 DB 325 VAIOORQVPGLEKEEESYVQAVMGAPRTCPBSMGCTNQQLRECHAFPTWMPDELGAISM 384
 QY 383 RVVYSYSAVVAVAHAHSLGCDKSTCKRVYIPWOLEETWKVNFLLDHOIFPDQ 442
 DB 385 SAAYNVYEAIVAAGHGLGCTSGTCARGVHMQLOQIYKVNFLHKKVAFDQK 444
 QY 443 DVALLLEIVQOMRSONPQFOSV--ASYVPLOROLKNIDISMHTVNNTPMSMKSRQ 500
 DB 445 DPLGIYDILANDMNGPEPTFEVIGSASLSPVHLDI-NKTKIOMHKNQOVAVCTRCL 503
 QY 501 SGOKRKPVGIHVCEFCIDCLPGFTLNTDEYECQACPNMWSYOSTSECFKROLVLE 560
 DB 504 EGHRLVYGHSCCEPCPCAGFLN-TSELHCOQPCGTEMAEAGSACFSRIVEELG 562
 QY 561 WHEAPTVALLAALGFLSTALIVIF-----WRHPTPLVNSAGSPMCLMTLLLV 613
 DB 563 WHE--PISLVLLAA--NTLLLLLLITAGLFAWR-LHTFVVRSAQRLFLMIGSLVA 615
 QY 614 AYVVVYVVGPPKYSTCLROALPFLCTICISIAVRSPOIYCAFKNASRPRAYSVW 673
 DB 616 GCSCLISFEGKPYVACCLLROPFLSLGPAFLSCLTIRSFQVILFEKSTVPFHYHWA 675
 QY 674 RYOGPVYMAFIVLKVIVYIGMLARPOSHPRDPP--DPKTIYVGCNPNYRSLFN 730
 DB 676 ONHGAGI--FVIVSYVHLFLCLLWLMMPRPFRYQORPHVLILCTEVSNGFLVA 732
 QY 731 TSLDILLVSGFRAVMAKELPTYNNEAKFTILSKTFYFTSSVSLCTMSAYSGVLVTV 790
 DB 733 FAHNILLISITFVCSYLKELPENYNEAKVTFSLHIFVSIAMIAFTMSSIYOGSYLPAV 792

QY 791 DLVTYVLLALLSLGYPKCYMILFPERNTPAVNSMIGQYMR 836
 DB 793 NVLAGLATLGGSGYFLPKCYVILCRPELNTNTEHFOASIDYTR 838

RESULT 10

ID Q8TDJ9 PRELIMINARY; PRT: 763 AA.

AC Q8TDJ9; 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, Last sequence update)
 DE 01-JUN-2002 (TREMblrel. 21, Last annotation update)
 DR Gml48 form B
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;

RP SEQUENCE FROM N.A.
 RX MEDLINE=21888635; PubMed=11891061;

RA Makalowska I., Good R., Faruque M.U., Hu P., Robbins C.M.,
 RA Eddings E.M., Mestre J.D., Baxevanis A.D., Carpten J.D.,
 RT Identification of six novel genes by experimental validation of
 RT GeneMachine predicted genes.
 RL Gene 284:203-213(2002)

DR EMBL; AF387618; AAL91359.1;
 SO SEQUENCE 763 AA; 84427 MW; 593616575D6BD17D CRC64;

Query Match 28.2%; Score 1253.5; DB 4; Length 763;
 Best Local Similarity 36.9%; Pred. No. 1.9e-95;
 Matches 287; Conservative 124; Mismatches 332; Indels 35; Gaps 12;

QY 75 MRFAVEEINNDSSLLPGVLLGEIVDVCYISNNQPVLYFLA---HEDNLLPIQEDYSN 130
 DB 1 MRLGVEEINNSTALLPNTLTGVLQVDCSANYATATRLVLSLPGQH---IELQGLIH 57
 QY 131 YISRVAVIYGPDSNSVTVANFLSLFLPQTYSAISDELNDKVRFPALLRTTSPADH 190
 DB 58 YSPYLVAVIGDSTIRAKTTAALLSPFLVPMIYASSETLSYKQVYSPFLRTINDKQ 117
 QY 191 VEAMVQMLHFRMNIIVLVSSDTYGRDNGQLLGERVARRDICIAFOETLPLQPNQNM 249
 DB 118 VETWVLLQKFGWMIISLVGSSDDYGLQVOALENQAAGGICIAFKDIMPFAQVGD-- 175
 QY 250 TSEKRLVTIVDKQOSTARVYVFPDLTLHFENEVLRFONFTGAVMIASEMAIDPV 309
 DB 176 ----ERMQCLMRHLAQAGATVYVFSRQALARVEFESVLTNLGKVMVASEAVALSRH 230
 QY 310 LHNLTGLHGLFTLIGITIOSVPIPGSEFSEFREMPOA--GPPPLSRTSOSTYCNQECNC 366
 DB 231 ITGVGIDIRIGVNLGVALQKRAVPGIKAFEEYARADKXARPCPKGSMC--SSNOLCRQC 289
 QY 367 LNATLSFNTLRLSGERVYVSYSAVVAVAHAHSLGCDKSTCKRVYIPWOLEETWK 426
 DB 290 QAFMAHTMPKLKAFSSMSAYNAVVAAGHGLHLLCASAGACSGRVYPMOLEQJHK 349
 QY 427 VNFITLDHOIFPDQDVALHLEIYQOMRSONPQFOSVAS--YTPLORLKNIDISM 484
 DB 350 VHFLLHKTVAFNDMRDPLSSYNIITAMWNGKMTFTYLGSTWSVPLDNI--NEPKIWH 408
 QY 485 TVNNTIPMSMKSRQOSQOKKPVGIHVCEFCIDCLPGFTLNTDEYECQACPNMWS 544
 DB 409 GKDNQVPSVSSDCLBEHQRYVTFHHCCEFCVCGAGTFLNKS-DLYKQPCGKEMA 467
 QY 545 YOSEISCFKROLVLEWHEAPTVALLAALGFLSTALIVLFWRHPTPLVNSAGSPMC 604
 DB 468 PEGSQTCPFRVYFLALNEHTSWVLLANTLLLLLLITAGLFAHNDLTPVRSAGRLC 527
 QY 605 FLMLTLLVATVAVYVYGPPKYSTCLROALPFLCTICISIAVRSFOIYCAFKNASR 664
 DB 528 FLMGSLAAGSGSLYGFEXEPFRACILROALFALGFIFPLSCLVRSFCLVRFSTK 587
 QY 665 FPRAYSVWVRQGPVYMAFIVLKVIVYIGMLARPOSHPRDPPDKITIVSCNPVYR 724


```

Db 588 VPEFYHAMVONHSGJLFVMISSAAQLICLTWLVWTPLPAREXORPPLHMECTET-- 645
QY 725 NSL-----FENTSLDLTSVGFSAFMGKELPTNYNEAKFTLSMTFVFTSSVSLCTE 778
Db 646 NSLGFITAFIYXNG-----LBSISAFACSYLTKDLPENYNEKCYFSLFNFVSIATFTT 701
QY 779 MSAYSGVATVTDLVTLVNLALISLGYFGPKCYMLFPERMTPAYNSMIGCYTMR 836
Db 702 ASYVDKYLPAANMMAGLSLSSGFGYFLPKCYVILCRDLNSTEHFQASIDQYTRR 759

RESULT 11
073635
ID 073635 PRELIMINARY; PRT; 940 AA.
AC 073635;
DT 01-AUG-1998 (TREMBLrel. 07, Created)
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Calcium2+ sensing receptor.
OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthopterygii; Acanthopterygii; Percomorphi; Tetraodontiformes;
OC Tetraodontidae; Takifugu.
OX NCBI_Taxid=31033;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98226788; PubMed=9560249;
RA Naito T., Saito Y., Yamamoto J., Nozaki Y., Tomura K., Hazama M.,
RT "Putative pheromone receptors related to the Ca2+-sensing receptor in
RT fugu."
RL Proc. Natl. Acad. Sci. U.S.A. 95:5178-5181(1998).
DR EMBL; AB008857; BAA26122.1; -.
DR InterPro: IPR001828; ANF_receptor.
DR InterPro: IPR000337; GPCR_Mgr.
DR InterPro: IPR000651; RasGEF.
DR Pfam; PF00003; 7tm_3; 1.
DR Pfam; PF01094; ANF_receptor; 1.
DR PRINTS; PR00248; GPCRMR.
DR PROSITE; PS00979; G_PROTEIN_RECPR_F3_1; 1.
DR PROSITE; PS0259; G_PROTEIN_RECPR_F3_4; 1.
KW Receptor.
SQ SEQUENCE 940 AA; 105814 MW; 06DAB7803B6878B3 CRC64;

Query Match 23.9%; Score 1062.5; DB 13; Length 940;
Best Local Similarity 30.1%; Pred. No. 1,9e-79;
Matches 263; Conservative 172; Mismatches 336; Indels 103; Gaps 23;

QY 16 WTAEPAENSDFTLPDGLLGSLFSLHANKG-----IVLNFLOVPMCKEY 62
Db 16 YVISTGPNORAOVTGILLGFLPIHFHGISKDNELANPESTKCVRFNR----- 67
QY 63 EVKVIQNLQARFAVEEINNDSSILPGVLGEIYDVC-YISNNQOPLYELA-HEDN 120
Db 68 -----GFRWQAVFALEEINNSSSLPNTLGRITDNTVYSKALEATLSFAQNKID 122
QY 121 LLLPIQE--DYSNTISRVAVIGPNSESVMVANFLSLFLPLQIYSAISDELDKVRFP 178
Db 123 SLNDEECNCTDHPATIAVGAAGSAVSTAVANLLSLFTIPQISVASSRLLSNKQYK 182
QY 179 ALKRTTSAHHYEAQMQLMHRMMWITLVSSDPTGRNGQLGVRVARDCICAFQE 238
Db 183 SFMKTIPTDHOATAMDAVEYFQMMNVIAVSDDDYGRGIEKFEKEMERDICHINE 242
QY 239 TLPTLPQNMWTSBERQLTVIVDKLOOSTARVYVSPDLTYHFENEVLKRONFGAVN 298
Db 243 LISQY-----FDCETALVDRIENSTAKYIVFASGPDLEPLIKEMVRNRITDRIW 294
QY 299 IASESNALDEVLNLTGLGTLGTTISVDPDIPGSEFERMGQAPPLSRTSQSYT 358
Db 295 IASEANASSSLIAKPEYLDVVEGIGVLAAGNIPGRFELQ---QVQPARGSINEVRE 351

```

```

QY 359 CNOECDNCL-----NATLSEFTILRLSGE-----RVVYSVYSAYV 393
Db 352 FWEETFCYLEDSPRIQESNGSDSPRLCTSEDLTSVETPYLDHNLHLSINRYAVY 411
QY 394 AVAAHLSLIG-----DKSTCKRVVYPMOLLEIKRVNFT-LDHQIFPDQGV 444
Db 412 SIHAALQDILICPPGHGLFANNSCADIKMEAOVLKQLHLVYTSMSGKRVHDEADM 471
QY 445 ALHEIYOWOMDSQN--PROSVASYPLQRO---LKNIDISMHVNNVTIMSMCKR 498
Db 472 EANTITNMHRSABDGVAFREVGYYHMAARKAKLLIDFTKMMWNNVSEVPNSCED 531
QY 499 COSGQKKRPV-GIHVCCFECDICLPGETFLNHTDEVECCAPNNMSYQSECTFKRLV 557
Db 532 CEPGRKGIIDSMPTCCFECECTECSDEYSDH-KDASICTCPNNMSNGHFTCLKEIE 590
QY 558 FLEWHEAPTAVALALAGLSTLAILIYWRHROPYIVASAGPQCFMLTLILVAVMY 617
Db 591 FLAMSEPPGIALAICAVGLALFAVGVVFRFRNPPIVASNRELSTYLLSLICFSS 650
QY 618 VPVYVGPVKSTCLCQALPLCFTTICISIAVRSPQIVCAFMAASRFPRA-YSYWRYQ 676
Db 651 SLIFIGEPQWTCRLKQPAFGISFVLCISILYKTRNLVLFV-AKIPSIHRKMGLN 708
QY 677 GPVYSAFIVLKVAVIVIGML-ARPOSHRTPDDPKRITIVSCNPYRNSLFFNTSLD 735
Db 709 LQFLVFLCTFVOVMICVWMLYNAPSSYRNHIDE--IIFICNBSGVALFLLGYTC 766
QY 736 LLSVGFSAFMGKELPTNYNEAKFTLSMTFVFTSSVSLCTMSAV---SGVLVTVLD 792
Db 767 LLAICFFFAFKSKRLPENFTENKFTFCMLIFFIYWI--FIPAFSTYGFVSAVEA 823
QY 793 LVTVLNLALISLG---YFGPKCYMLFPERMT 822
Db 824 IA-----ILASSYGLACIFPNKYIILFKRCRMT 853

RESULT 12
0923K1
ID 0923K1 PRELIMINARY; PRT; 858 AA.
AC 0923K1;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Sweet taste receptor T1R3.
GN TAS1R3.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_Taxid=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MISTAR;
RA Nelson G., Hoon M.A., Chandrasekar J., Zhang Y., Ryba N.J.P.,
RA Zuker C.S.,
RT "Mammalian Sweet Taste Receptors."
RL Cell 0:0-0(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MISTAR; TISSUE=CIRCUMVALLATE PAPILLA;
RX MEDLINE=21927605; PubMed=11917125;
RL X., Staszewski L., Xu H., Dutler K., Zoller M., Adler E.;
RT "Human Receptors for Sweet and Umami Taste."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4692-4696(2002).
DR EMBL; AY032620; AAK51601.1; -.
DR EMBL; AF456324; AAM10636.1; -.
DR InterPro: IPR001828; ANF_receptor.
DR InterPro: IPR000337; GPCR_Mgr.
DR Pfam; PF00003; 7tm_3; 1.
DR Pfam; PF01094; ANF_receptor; 1.
DR PROSITE; PS00980; G_PROTEIN_RECPR_F3_2; UNKNOWN_1.
DR PROSITE; PS0259; G_PROTEIN_RECPR_F3_4; 1.

```



```

QY 604 CEMLTLLVAVVPPVYGPVKVSTCLCQALPFLCTICISCIARSPQIVCAFMMAS 663
DB 636 SYVLLFLICFESSSLIFIGQPODMCRLQRPARGISFVICISILVKTNNVLLVFE--A 693
QY 664 RFPRA-ISTVWRVGPVYSMAFTVLMKVIVIGML-ARQSHRTPDPDKITIVSCNP 721
DB 694 KIPSLIRKRWGMLNGLLFLVLCFVQVMICVWVLYNAPSSYNHIDE--IIFITCNE 751
QY 722 NYRSLFNTSLDLLSVGSPFAYMGKELPTVYNEAKFTITLSMTFTYSVSLCTFMSA 781
DB 752 GSWALFFLLGYTCLLAICFFFAFKSRKLPENFTAKFTFESMLIFFIYWTIS--TIPA 808
QY 782 Y---SGVLTIVDLVLTVMNLALSLG----YFGPKCYMILFYPERNT 822
DB 809 YFSYTGKFSVAEVAIA----ILASSFGMLACIFENKVIYILFKPSRMT 852

RESULT 14
Q925D8 PRELIMINARY; PRT; 858 AA.
AC Q925D8;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative sweet taste receptor family 1 member 3.
GN TAS1R3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SVEV;
RX MEDLINE=2125287; Pubmed=11326277;
RA Max M., Shanker Y.G., Huang L., Rong M., Campagne F., Liu Z.,
RA Weinstein H., Damak S., Margolskee R.F.;
RT "Tas1r3, encoding a new candidate taste receptor, is allelic to the
RT sweet responsiveness locus Sac.";
RL Nat. Genet. 28:58-63(2001).
DR EMBL; AF368025; AAK5537.1; -.
DR MGI; MGI:1933547; Tas1r3.
DR InterPro; IPR001828; ANF_receptor.
DR InterPro; IPR000337; GPCR_Mgt.
DR Pfam; PF00003; 7tm_3; 1.
DR Pfam; PF01094; ANF_receptor; 1.
DR PROSITE; PS00980; G_PROTEIN_RECP_F3_2; UNKNOWN_1.
DR PROSITE; PS50259; G_PROTEIN_RECP_F3_4; 1.
KW Receptor.
SQ SEQUENCE 858 AA; 94533 MW; 2012102B55D7E6B0 CRC64;

Query Match 22.9%; Score 1018; DB 11; Length 858;
Best Local Similarity 29.7%; Pred. No. 8,6e-76;
Matches 251; Conservative 164; Mismatches 351; Indels 80; Gaps 22;

QY 24 NSDFYLPDYLGLGFLSLANMKGIHNLQV--MCKREYKVIQYINLQAMRRAYE 80
DB 26 SQCFKAGDYITGLFLPLGSTEAT--LNORAQNSTLCNFF--SPGLFLAMAKAYE 81
QY 81 EINDSSLLGVLGELIVVCYISNNVQY-----LYFLAH-EDNLLPQEDYSNTIS 133
DB 82 EINGSSLLPGLRLGYLDFTC-----SEPVYMKSSIMFLAKVGSOSIAYCVTYQOP 136
QY 134 RVAVVVPDNSSEYMTANFLSLFLPQITYAISDELQKVRPALRTTPSADHVEA 193
DB 137 RVLAIVGPHSSELALITGKFFSFLMPQVSTASMDRLSDRETPPSFLRTVPSDVLQA 196
QY 194 MVOLMLFRNNWITIVVSDPYGRNQQLGERVARDICIAFOETLPTIQPNQNMTEE 253
DB 197 VVTLQVSNWVVAALGSDDDYGRGLSIFSSLANARGICIAHE---GLVPQHDTSGQO 252
QY 254 KQRLVTVDKLQOSTARVAVVSPDLTLVHFENEVLRQNTFGAVWIASAIDPVLAHL 313

```

```

DB 253 LGRVLDVLRQVNSKVOVVVLFASARAAYSLFSYIHHGLSPKRWVASESMTSLDVLMTL 312
QY 314 TELGHLGTFPIGITIOSVPIGSEFREMGPQAGPP-----LSRTQSYTCNOE 362
DB 313 PNIAVGTIVGLQRLGALLPEFSHYETHALADPAFCASINAEDELREHWGQRCPO- 371
QY 363 CDNCLNNTLSFNTILRISGERV---VYSVYSAVYAAHALSLGCDKSTC-TRVYVYV 418
DB 372 CDDIMLQNLSSGLLQNLNSAGOLHHQIPATYAAYVSAQALHNTLQCNVSHCHSEVLPW 431
QY 419 QLEELKRVNFTLDHIFEDPDQGVVALHLEIYQWMDRQNPQVASTYPIQRLQKNI 478
DB 432 QLENNMYMSFHARDLQLQDABGNVMEYDLKMMVQSPPTVLAHTVGT-NGTLQ 487
QY 479 QDISMHTVNTTIPMSKSKQSGOKKPVGIHVCEPCIDCLPGFLNHTDEDEYCOAC 538
DB 488 QQSKMYRGNQVPSQCSROCKQGVARVGFHSCCYDCDCAKAGSYRKH-PDFTCTPC 546
QY 539 PNNEMSYQSETSCFKROLVLEWHEDPTIAVALLALGLSTALIVFRHRQTPIVRS 598
DB 547 NODQMSPEKSTACLPRRPKFLAMGEPPVLLILLICLVGLALALALSYHHMDSPLVQA 606
QY 599 AGG-PNCFMLTLLVAVVPPVYGPVKVSTCLCQALPFLCTICISCIARSPQIVC 657
DB 607 SGGSQFCFGLICLGLFCLSVL-LFPGRPSSASCLAQPMHMLPLTGCLSTFLQAAETP- 664
QY 658 AFKMASRFPRAYSY---RYQGPVYSMAFTVLMKVIVIGMLARPOSHRTPDP-- 710
DB 665 ---VESLPLSMANWLCISYRLGAMLVLLATFVEAALCAMLTAAP---PEVYDMSV 718
QY 711 DPKTIVSCNPNRNSLFTNTSLDLLSVGFSRNVKELPTVNEAKFTITLSMTFTFT 770
DB 719 LPTFVLEHCHVRVSWSLGVLHITNAMLAFCLFTPLVQSGPGRYNRRARGITFAMLAFT 778
QY 771 SSVS---LCTFMSAY-----SGVLTIVDLVLTVMNLALSLGYPGKCYMILFYPER 820
DB 779 TWVSFVLLANVOVAIOPAVOMGAILVCAIGIYV-----FHLPCYVILMLPKL 828
QY 821 NTPAYF 826
DB 829 NTQGEF 834

RESULT 15
Q925A4 PRELIMINARY; PRT; 858 AA.
ID Q925A4;
AC Q925A4;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative taste receptor.
GN TAS1R3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RA Saint E., Koriely J.N., Battey J.F., Sullivan S.L.;
RT "Identification of a novel member of the T1R family of putative taste
RT receptors.";
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AV026318; AAK01937.1; -.
DR InterPro; IPR001828; ANF_receptor.
DR InterPro; IPR000337; GPCR_Mgt.
DR Pfam; PF00003; 7tm_3; 1.
DR Pfam; PF01094; ANF_receptor; 1.
DR PROSITE; PS00980; G_PROTEIN_RECP_F3_2; UNKNOWN_1.
DR PROSITE; PS50259; G_PROTEIN_RECP_F3_4; 1.
KW Receptor.
SQ SEQUENCE 858 AA; 94567 MW; A7879C015C0E8474 CRC64;

```

Query Match	22.8%;	Score 1014;	DB 11;	Length 858;
Best Local Similarity	29.6%;	Pred. NO. 1.8e-75;		
Matches 250;	Conservative 164;	Mismatches 352;	Indels 80;	Gaps 22;

QY	24	NSDPEYLPDYLGGELFSLSHAMKQITVHLEQVY-----MCKEYEVKIVGYNLMQAMRPAVE	80
Db	26	SOQKRAQGDYTLGGELFPLGSTEET--LNOAOPNSTILCRF--SPYGLFLAMAMKAYE	81
QY	81	EINDDSLPLGCVLLGEYELVDYCISSNNQVPY-----LYFLAH-EDNLIPLIOEDYSNYIS	133
Db	82	EINNSGALLPPLRIGYDLDFTC-----SEPVTWKSSILMAKVGSQSIYAACNYQYOP	136
QY	134	RVAVAVIGDNESEWTVYANFLSLELLDQIYTSAISDELPRKVRPALLRTPPSADHVEA	193
Db	137	RYLAATIGHSESELALTIGKEFFSEFLMPOVYSASMDRISIRETFPPSEFRTPVSPDRVOLA	196
QY	194	MYOLMIFRRMMIIVYSSDPYGGDNOLLGERVARDICIAFOETLPTLOPQNMNMTSE	253
Db	197	VYTLLOQNSWMMVAAALGSDDDYDGEGLSTSPSSLANAGICIAHE-----GLYPOHDISGOQ	252
QY	254	RORLVTIYDKLQOSTARVAVVSPDLTYHFEFNEVLNQNFTGAVIASESMAIDPVLHNL	313
Db	253	LKATLYDLRYQVNOBKQVYVYLFASARAVYLSFYSIHGGLSPKVVWASEMLTSDLMWTL	312
QY	314	TELGHLGTFELGIIIOVPIPFSEFRMRQAPRP-----LSTQSYTCNOE	362
Db	313	PNIARVGVIVLGFLORGALLPFEFSHYVETHALADRPFCASLNAEILDENHWQORPO-	371
QY	363	CDNCNLATLSFNTILRLSGERY--VYSVYSAYVAVAAHALHSLGCDKSTC-TRRVYPM	418
Db	372	CDDIMLQNLSSGLLONTSAGOLHQQIFATYAAVYSVAOLHNTLQCNVSHCHEVHLPM	431
QY	419	QLLEIMKVNVTLLDQOIEPRPOGDVALHLEIYQOMQDRSQNPQSVASYRPLQORLKN	478
Db	432	QLLENNYKMSHADLTLLQFPABENVDMEYDKMMWQOSPTPVLYHTVGT-----NGTLQ	487
QY	479	QDISMHTVNNNTIPMSMCKRQOSQOKKRPVGIHVCSCECIDCLPRTGLNHTHEDEYEOAC	538
Db	488	QOSKMYMGQNVQPVQSCROCKDQGVRRVKGHSCCYDVCYDKAGSYRKH-PDFTCTPC	546
QY	539	PNNEMSVQSEIICKROLVLEMEHAEPTIVALLAAGLSTALIVYFMRKHQTPYRS	598
Db	547	NODQMSPEKSTACLPRRRPKFLAMEBPVLSILLLLCLVLGIALAALGLSVHNHDSPLYQA	606
QY	599	AGG-PMCFIATLTLVAVMYVPVYVGPBKYSTCICROALRPLCTICISIAVRSQIYC	657
Db	607	SGGSGOFCGGLCGLFELCLSVL-LEPRGPRSSACIQAOPPAHLPLTGLISTLFLQAAETP-	664
QY	658	AFKMAAREPRAYSYW---VRQGPVYSMAFTVYLKMYIVYIGMLARPOSHRPTDP--	710
Db	665	---VESELPISLMANMLCSYRLGMLAMLVYLLATFEVALCAWYLTPAR---PEVYDMSV	718
QY	711	DPKTTIYSCNPNYNSLILFNSTLIDLILSVYGFSPRAYGKELPITNYNNAKITYLSMTFFY	770
Db	719	LPTLEVLEHCHRRSVNSLGLVHTTNAMIAELCFGLTFLVQSOGRYNNARARILTPAMLAFT	778
QY	771	SSVS-----LCFMSAY-----SGVLVTIYDLVATVYNLNLASIGYGRPCYUMLIPEP	820
Db	779	TWVSFVPLLANVQVAYQPAVQMGAILVICALGILT-----FILPRCYVLLMLPKL	828
QY	821	NTPAYF	826
Db	829	NTQEFF	834

```
Search completed: May 19, 2003, 09:52:37
Job time : 43.1811 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:48:40 ; Search time 14.3799 Seconds
(without alignments)
1714.646 Million cell updates/sec

Title: US-09-927-315-9
Perfect score: 4443
Sequence: 1 MGPRAKTICSLFLMLVLA.....ERNTPAYENMIGYTRRD 838

Scoring table:
BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*

- 1: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
- 2: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
- 3: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
- 4: /cgn2_6/ptodata/1/1aa/6B.COMB.pep:*
- 5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
- 6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3514	79.1	669	4	US-09-361-631-7 Sequence 7, Appli
2	3203.5	72.1	843	4	US-09-361-631-1 Sequence 1, Appli
3	3151.5	70.9	843	4	US-09-361-631-2 Sequence 2, Appli
4	1107	24.9	1059	4	US-09-134-513-2 Sequence 2, Appli
5	1094	24.6	1078	1	US-08-485-588-7 Sequence 7, Appli
6	1094	24.6	1078	1	US-08-484-565-7 Sequence 7, Appli
7	1094	24.6	1078	2	US-08-480-751-7 Sequence 7, Appli
8	1094	24.6	1078	2	US-08-943-966-7 Sequence 7, Appli
9	1094	24.6	1078	3	US-08-353-784-7 Sequence 7, Appli
10	1094	24.6	1078	3	US-08-484-719B-7 Sequence 7, Appli
11	1094	24.6	1078	4	US-08-484-159-7 Sequence 7, Appli
12	1092.5	24.6	1027	4	US-09-162-021B-2 Sequence 7, Appli
13	1088.5	24.5	1079	1	US-08-485-588-8 Sequence 8, Appli
14	1088.5	24.5	1079	1	US-08-484-565-8 Sequence 8, Appli
15	1088.5	24.5	1079	1	US-08-480-751-8 Sequence 8, Appli
16	1088.5	24.5	1079	2	US-08-943-966-8 Sequence 8, Appli
17	1088.5	24.5	1079	3	US-08-353-784-8 Sequence 8, Appli
18	1088.5	24.5	1079	3	US-08-484-719B-8 Sequence 8, Appli
19	1088.5	24.5	1079	4	US-08-484-159-8 Sequence 8, Appli
20	1087.5	24.5	1085	1	US-08-485-388-5 Sequence 8, Appli
21	1087.5	24.5	1085	1	US-08-484-565-5 Sequence 5, Appli
22	1087.5	24.5	1085	2	US-08-480-751-5 Sequence 5, Appli
23	1087.5	24.5	1085	2	US-08-943-966-5 Sequence 5, Appli
24	1087.5	24.5	1085	3	US-08-353-784-5 Sequence 5, Appli
25	1087.5	24.5	1085	3	US-08-484-719B-5 Sequence 5, Appli
26	1087.5	24.5	1085	4	US-08-484-159-5 Sequence 5, Appli
27	1079	24.3	1088	4	US-08-485-588-6 Sequence 6, Appli

28	1079	24.3	1088	1	US-08-484-565-6 Sequence 6, Appli
29	1079	24.3	1088	2	US-08-480-751-6 Sequence 6, Appli
30	1079	24.3	1088	2	US-08-943-966-6 Sequence 6, Appli
31	1079	24.3	1088	3	US-08-353-784-6 Sequence 6, Appli
32	1079	24.3	1088	3	US-08-484-719B-6 Sequence 6, Appli
33	1079	24.3	1088	4	US-08-484-159-6 Sequence 6, Appli
34	1012.5	22.8	877	4	US-09-619-353-2 Sequence 2, Appli
35	995	22.4	1219	2	US-08-687-289A-6 Sequence 6, Appli
36	917.5	20.7	863	4	US-09-619-353-14 Sequence 14, Appli
37	867.5	19.5	851	4	US-09-619-353-12 Sequence 12, Appli
38	856	19.3	854	4	US-09-619-353-10 Sequence 10, Appli
39	850	19.1	856	4	US-09-619-353-8 Sequence 8, Appli
40	849.5	19.1	835	4	US-09-619-353-7 Sequence 7, Appli
41	743	16.7	1058	2	US-08-687-289A-5 Sequence 5, Appli
42	724.5	16.3	915	2	US-08-453-862-2 Sequence 2, Appli
43	724.5	16.3	915	1	US-08-452-734A-2 Sequence 2, Appli
44	724.5	16.3	915	4	US-08-176-401B-2 Sequence 2, Appli
45	724.5	16.3	915	5	PCT-US94-14989-2 Sequence 2, Appli

ALIGNMENTS

```
RESULT 1
US-09-361-631-7
; Sequence 7, Application US/09361631
; Patent No. 6383778
;
GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Lindemeier, Juergen
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
; TITLE OF INVENTION: Involved in Sensory Transduction
; FILE REFERENCE: 02307E-08672005
; CURRENT APPLICATION NUMBER: US/09/361,631
; EARLIER FILING DATE: 1999-07-27
; EARLIER APPLICATION NUMBER: US 60/095,464
; EARLIER FILING DATE: 1998-07-28
; EARLIER APPLICATION NUMBER: US 60/112,747
; EARLIER FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
;
LENGTH: 669
;
TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human G-protein coupled receptor (GPCR) B4 amino
; OTHER INFORMATION: acid sequence
US-09-361-631-7
Query Match 79.1%; Score 3514; DB 4; Length 669;
Best Local Similarity 98.4%; Pred. No. 0;
Matches 666; Conservative 1; Mismatches 2; Indels 8; Gaps 2;
```

QY	162	ITVSAISDELARKVRRPALFTTPSADHYEAMQMLHFRNNWITYSSPTGRRNQ	221
DB	1	ITVSAISDELARKVRRPALFTTPSADHYEAMQMLHFRNNWITYSSPTGRRNQ	60
QY	222	ILGSEVARRDICIARQETPLTLPQPNOMTSEERORLVTIVDLQOSTARVVVFPDILL	281
DB	61	ILGSEVARRDICIARQETPLTLPQPNOMTSEERORLVTIVDLQOSTARVVVFPDILL	120
QY	282	YHFEVNEVLRONTTGVAVWIASSEWALDPVLAHNTLGLGCTGTTQSVPIPGSEFRW	341
DB	121	YHFEVNEVLRONTTGVAVWIASSEWALDPVLAHNTLGLGCTGTTQSVPIPGSEFRW	180
QY	342	GQOAPPPPLSRFSOSYTQOECDCNCLNATLSFNTILRLSGEVVYVSVAVAVAHALHS	401
DB	181	GQOAPPPPLSRFSOSYTQOECDCNCLNATLSFNTILRLSGEVVYVSVAVAVAHALHS	240
QY	402	ILGCDKSTCTKRVVYPMWLLEIRVNFLLDHOIFPDQGVALLHLIVQMWRQNP	461

```

Db 241 LLDGDKSTCKRKYVYPOLLEIEIKVNFETLDDHGFDPQGDVALHLEIVQMOMDRSQNP
QY 462 FQSAVASYPIOROLAKNIODISMTHTVNNITIPMSMCKRQSGOKKKPVGIHVCCFECIDCL
Db 301 FQSAVASYPIOROLAKNIK-75LHTVNNITIPMSMCKRQSGOKKKPVGIHVCCFECIDCL
QY 522 PGFELNTEDEYECQACPNMWSYQSETSCFRKQVLFEWHEAPTLAVALLAAGFLSTL
Db 360 PGFELNTEDEYECQACPNMWSYQSETSCFRKQVLFEWHEAPTLAVALLAAGFLSTL
QY 582 AIIYIFRHRQPTIVRSAGGPMCFMLTLVLAAMVVPYVGPVKVSCICLROLAPLFCF
Db 413 AIIYIFRHRQPTIVRSAGGPMCFMLTLVLAAMVVPYVGPVKVSCICLROLAPLFCF
QY 642 TICISCLAVRSFOIVCAFKNASRPRAVSYVWRQGPVSNAPITVLKMYIVIGMLARP
Db 473 TICISCLAVRSFOIVCAFKNASRPRAVSYVWRQGPVSNAPITVLKMYIVIGMLARP
QY 702 QSHRTPDDPKITIVSCNPNYRNSLLENTSLDLLSVGFSFAYMGRKELPTNNEAKFI
Db 533 QSHRTPDDPKITIVSCNPNYRNSLLENTSLDLLSVGFSFAYMGRKELPTNNEAKFI
QY 762 TLMSTFEFTSSVSLCTFMSASGVLTIVDLTVVNLALSLGYPGPKCMILFPERN
Db 593 TLMSTFEFTSSVSLCTFMSASGVLTIVDLTVVNLALSLGYPGPKCMILFPERN
QY 822 TPAYFNSMIGYTMARD 838
Db 653 TPAYFNSMIGYTMARD 669

```

RESULT 2

```

US-09-361-631-1
; Sequence 1, Application US/09361631
; Patent No. 6383778
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliott
; APPLICANT: Lindemeyer, Juergen
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids Encoding a G-protein Coupled Receptor
; FILE REFERENCE: 02307E-08872005
; CURRENT APPLICATION NUMBER: US/09/361,631
; EARLIER FILING DATE: 1999-07-27
; EARLIER APPLICATION NUMBER: US 60/095,464
; EARLIER FILING DATE: 1998-07-28
; EARLIER APPLICATION NUMBER: US 60/112,747
; EARLIER FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentln Ver. 2.0
; SEQ ID NO 1
; LENGTH: 843
; TYPE: PRT
; ORGANISM: Rattus sp.
; FEATURE:
; OTHER INFORMATION: rat G-protein coupled receptor (GPCR) B4 amino
; OTHER INFORMATION: acid sequence
US-09-361-631-1

```

```

Query Match 72.1%; Score 3203.5; DB 4; Length 843;
Best Local Similarity 70.6%; Pred. No. 7.5e-296;
Matches 595; Conservative 108; Mismatches 133; Indels 7; Caps 4;

```

```

QY 1 MGPRAKTCISFELLAWLAEP--AENSDFPLDGDYLLGFGFSHAMKGVHNFQVY 57
Db 1 MGPARTLCILSLHLPLKPKGLVENSDFLLAGDYLLGGLFTLHANKVSIHSLSYLQVP 60
QY 58 MCKEYEVKVGYNLMQAMRAVEEINNDSSLLPGVLGLEYVDVGCYSNNQVPLVLAH 117
Db 61 KCNFTMKVGLGYNLMQAMRAVEEINNDSSLLPGVLGLEYVDVGCYSNNQVPLVLAH 120

```

```

QY 118 EDNLPIQIEDYSNTISRYAVVIGPDNSESVMYANFSLPILPOITYSAISDELDRKYR 177
Db 121 DDDLPLIKDKYSQYMPHVAVVIGPDNSESATIVSNIIISHELPQITYSASIDKLRDRHF 180
QY 178 PALIRTPSADHVEAVOVLHFRMRMTIYVSSDPRYGRDNQOLGERRAR-RDICIATF 236
Db 181 PSMILRTVPSATHTHEAVOVLHFRMRMTIYVSSDPRYGRDNQOLGERRAR-RDICIATF 240
QY 237 QETPLTQPNOMNTSEERORLVTVDRLOOSTARVAVVSPDLTHAFNEVLRONFTGA 296
Db 241 QEVLPFIPESQVMSRQROLDNLKLRISAVVAVVSPDLTHAFNEVLRONFTGA 300
QY 297 VWIASESWAIDPVLANHTEGLCTFLGCTIYQVPIPGSEFERMGPOACBPPLISRTSOS 356
Db 301 VWIASESWAIDPVLANHTEGLCTFLGCTIYQVPIPGSEFERMGPOACBPPLISRTSOS 360
QY 357 YTCNOECNCLNATLSNTLRLSGEVRVYSVAVVAHAHLSLGCOKSTCKRVVY 416
Db 361 TTCMOCDCLNTKSNLNLISGERVYSVAVVAHAHLSLGCOKSTCKRVVY 420
QY 417 PMQLLEEIMKVNFTLDDHGFDPQGDVALHLEIVQMOMDRSQNPQFQSAVSYPIOROLK 476
Db 421 PMQLLEEIMKVNFTLDDHGFDPQGDVALHLEIVQMOMDRSQNPQFQSAVSYPIOROLK 480
QY 477 NIQDISMTHTVNNITIPMSMCKRQSGOKKKPVGIHVCCFECIDCLPOTFLNTEDEYECQ 536
Db 481 YINNVSWTTPNNTIPVMSCKRQSGOKKKPVGIHVCCFECIDCLPOTFLNTEDEYECQ 540
QY 537 ACPPNEMWSYQSETSCFRKQVLFEWHEAPTLAVALLAAGFLSTLALIVFMRHRQPTIV 596
Db 541 SCPPNEMWSYQSETSCFRKQVLFEWHEAPTLAVALLAAGFLSTLALIVFMRHRQPTIV 600
QY 597 RSAGGPMCFMLTLVLAAMVVPYVGPVKVSCICLROLAPLFCF TICISCLAVRSFOIV 656
Db 601 RSAGGPMCFMLTLVLAAMVVPYVGPVKVSCICLROLAPLFCF TICISCLAVRSFOIV 660
QY 657 CAFTMASRFPRAVSYVWRQGPVSNAPITVLKMYIVIGMLARPQSH-RTPDDPKI 714
Db 661 CAFTMASRFPRAVSYVWRQGPVSNAPITVLKMYIVIGMLARPQSH-RTPDDPKI 719
QY 715 TIVSCNPNYRNSLLENTSLDLLSVGFSFAYMGRKELPTNNEAKFTLSTFEFTSSVS 774
Db 720 TIVSCNPNYRNSLLENTSLDLLSVGFSFAYMGRKELPTNNEAKFTLSTFEFTSSVS 779
QY 775 LCTFMSASGVLTIVDLTVVNLALSLGYPGPKCMILFPERNTPAYFNSMIGYT 834
Db 780 LCTFMSASGVLTIVDLTVVNLALSLGYPGPKCMILFPERNTPAYFNSMIGYT 839
QY 835 MRR 837
Db 840 MRR 842

```

RESULT 3

```

US-09-361-631-2
; Sequence 2, Application US/09361631
; Patent No. 6383778
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliott
; APPLICANT: Lindemeyer, Juergen
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids Encoding a G-protein Coupled Receptor
; FILE REFERENCE: 02307E-08872005
; CURRENT APPLICATION NUMBER: US/09/361,631
; EARLIER FILING DATE: 1999-07-27
; EARLIER APPLICATION NUMBER: US 60/095,464
; EARLIER FILING DATE: 1998-07-28
; EARLIER APPLICATION NUMBER: US 60/112,747
; EARLIER FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patentln Ver. 2.0

```



```

; SEQ ID NO 2
; LENGTH: 843
; TYPE: PRT
; ORGANISM: Mus sp.
; FEATURE:
; OTHER INFORMATION: mouse G-protein coupled receptor (GPCR) B4 amino
; OTHER INFORMATION: acid sequence
US-09-361-631-2

```

```

Query Match      70.9%; Score 3151.5; DB 4; Length 843;
Best Local Similarity 68.9%; Pred. No. 6.6e-291;
Matches 581; Conservative 113; Mismatches 142; Indels 7; Gaps 4;

```

```

OY 1 MGPRAKTCISLFFLLMVALEP---AENSDFYLGSDYLGLFSLHANKGIYHNFLOQP 57
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 1 MGPQATHTLHLLFLHLPKPMALVGNSEDFLAGDYLGGLETLHANAKSVSHLSLOYLP 60
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
OY 58 MCKEYEVKIVGYNLMQAMFAVEEINNDSSLPGVLLGEYDVGYISNNVQPVLEFLAH 117
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 61 KCNEYNMKVLGYNLQAMFAVEEINNCSLLPGVLLGEYDVGYISNNVQPVLEFLAH 120
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
OY 118 ENLLPIQEDISNTISRYAVYIGPDNSVMTVANFLSLFLPQITYSAISDELDRKARF 177
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 121 IDDFPLIKDYQYRPOVAVYIGPDNSSESAITVSNILSYFLVPOVYSAITDKLQDRKF 180
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
OY 178 PALFTTPSADHHEAMVOLMHEFRNMWITIVSSDPTGRODGOGLGEVARR-DICTAF 236
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 181 PAMKTRVPSATHIEAMVOLMHEFRNMWITIVSSDPTGRODGOGLGEVARR-DICTAF 240
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
OY 237 QETLPLQPNOMTSEERQRLVTVDKLQOSTARVVVVFSPDLTYHFEVEVLRONFTGA 296
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 241 QEVLPVPEPNOAVRPEEOQLNIIDKLRTSARVVVVFSPDLTYHFEVEVLRONFTGF 300
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
OY 297 VWIASPSAIDVVLNLTFLGHLGTELTITQSVPIPGSEFRMGPOGPPPLSTQSS 356
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 301 VWIASPSAIDVVLNLTFLGHLGTELTITQSVPIPGSEFRMGPOGPPPLSTQSS 360
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
OY 357 YTCNOCNDCLNATLSFTILRLGSEVVYVYSAVYAVALHSLGDCSTCRKRYV 416
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 361 TTCNDQDCAQNMITESEFNVLMLSEGRVYVYSAVYAVALHSLGDCSTCRKRYV 420
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
OY 417 PMQLLEETKVNFTLLDHOIFPDQDVALHLEIVQMDRSONPQVASYYPPLQRLQK 476
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 421 PMQLLEETKVNFTLLDHOIFPDQDVALHLEIVQMDRSONPQVASYYPPLQRLQK 480
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
OY 477 NIQDISMHTVNTIMSMKSCRCQSGOKKRGVGHVCCPECIDCLPFTFLNTEDEYEQ 536
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 481 YISNVSWTPNNTVPIISMCKSCQPGQMKPIGLHPCCEECVDCPDYILNRSVDEFNCL 540
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
OY 537 ACPNEMWSQSESECFKQVLVLEWHEAFTIYVALLAALGFSTALITVIFRHHQTPV 596
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 541 SCFGSMWSKNNIACFKRLAFLHWHVPTIYVITIAALGFSTALITVIFRHHQTPV 600
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
OY 597 RSAGSPMCFMLTLLLVAVVYVYVGPVKSTCLCQALFPLCFTTICISIAVSFOIV 656
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 601 RSAGSPMCFMLTLLLVAVVYVYVGPVKSTCLCQALFPLCFTTICISIAVSFOIV 660
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
OY 657 CAFKASRPRAYSYWVRQGPVYMAFTVYKMTIVYIGMARQSHR--TDDDDPKI 714
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 661 CAFKASRPRAYSYWVRQGPVYMAFTVYKMTIVYIGMARQSHR--TDDDDPKI 719
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
OY 715 TVISGNPNRNSLFTNTSIDLLSVVGFSAVYKELPTNYNEAKFTLSMTFTSSVS 774
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 720 IILSCHPNRNGLLNTSIDLLSVVGFSAVYKELPTNYNEAKFTLSMTFTSSVS 779
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
OY 775 LCTFSANSGVLVTVLLVTVNLIALISLGYFGPKCYMILFYPERNNPAYNSMIOGYT 834
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 780 LCTFSANSGVLVTVLLVTVNLIALISLGYFGPKCYMILFYPERNNPAYNSMIOGYT 839
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
OY 835 MRR 837
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
DB 840 MRR 842
| | | | | : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

```

```

RESULT 4
US-09-134-513-2
; Sequence 2, Application US/09134513
; Patent No. 6210964
; GENERAL INFORMATION:
; APPLICANT: Brown, Edward M.
; APPLICANT: Diaz, Ruben
; APPLICANT: Bai, Mei
; TITLE OF INVENTION: The Avian Extracellular Calcium-Sensing
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Vinson & Elkins L.L.P.
; STREET: 1455 Pennsylvania Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.
; ZIP: 20004-1008
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/134,513
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sanzo, Michael A.
; REGISTRATION NUMBER: 36,912
; REFERENCE/DOCKET NUMBER: BR1331/13003
; TELEPHONE: (202)639-6585
; TELEFAX: (202)639-6604
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1059 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: protein
; HYPOTHEICAL: NO
US-09-134-513-2

Query Match      24.9%; Score 1107; DB 4; Length 1059;
Best Local Similarity 31.9%; Pred. No. 3.8e-96;
Matches 284; Conservative 148; Mismatches 354; Indels 104; Gaps 24;

```

```

QY 358 TCNOCDNCL-----NATLSEN-----374
: | |
: | |
Db 350 EMEETFCNCLPSEKSNPASAFHKAHEBGLGANGNAARPCOTGENTSVETPYMD 409
: | |
QY 375 -TILRIGERVYVYSAVAVAHLSLGC-----DKSTCRVYVPMOLLEETW 425
: | |
: | |
Db 410 FTHLRIS-----YVYVAVASIAHALODIYCTPGKGLFTNGSCADIKKAVAMOVKLRLR 464
: | |
QY 426 KVNFTL-LDHOIFPDPOGDVALHLEIYOMOMDRSONP--FOSVASYPL-----OROLKNI 478
: | |
: | |
Db 465 HINTSMNGEYVDDEFGDLVGNSTIIMHLSPEBGSVEEVEGHVYNAKKELEFINE 524
: | |
QY 479 ODISMHTVNNITPMSCKRCQSQOKKPV-GIHVCECEIDCLPGFLNTHTEDEYECOA 537
: | |
: | |
Db 525 NKILMSGSKVEVFPNSCRDCLPGRKGIIEGPTCECEYDCPDGEYSDET--DASACK 583
: | |
QY 538 CPNNEMVYQSTSCFKQVLVLEHNEAPTAVALLAALGFLSTALIVIFRHHQTPYVR 597
: | |
: | |
Db 584 CPEDYWSNENHSCIPQIEFLSWEPEGLALTFEAVLGIFLTSFVLGVFKFRNTPIVK 643
: | |
QY 598 SAGPMCFMLTLLVAVYVGPVKVSTCLROALPCLPCLTICISIAVRSFOIYC 657
: | |
: | |
Db 644 AFNHELSTLILFSLICPSSSLFTIGEPQNTCHLRQPAFISVYLCISCLIVKTRVYL 703
: | |
QY 658 AFKMASRPPRA-YSYWARYQPVYSMAFTVLMKVIVIGMLARPOSHRTDPPDKITI 716
: | |
: | |
Db 704 VFE--AKIPTSLHRRMGLNQLFLVLCFTVQIVICVIMVYTAAPSRYNHELEDEITE 761
: | |
QY 717 VSCPNRNLSLLENTSLDLISVGFSPAYMGKELPTNYNEAKITTSMTFFSVSLC 776
: | |
: | |
Db 762 ITCHGSLMAGFLIGYTCALACIFFEAFKSRKLPENFNKAKFTLSMLTFFIWMIS-- 819
: | |
QY 777 TFMASGCVLTIVDLTVYNLALISG---YEGPKCYMLTPEPENT 822
: | |
: | |
Db 820 -FIRAYASTYKRVY-AVEYAIILASPGILACIFENKVIYILKRPKSNPT 867
: | |

```

```

APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1078 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-588-7
Query Match 24.6%; Score 1094; DB 1; Length 1078;
Best Local Similarity 31.5%; Pred. 6.7e-95;
Matches 279; Conservative 160; Mismatches 342; Indels 106; Gaps 26;
QY 16 WTAEAPENSDFYLP-----GDYLLGLFSLHANKGIYHNLQVPMCKEYEVKYG 68
: | |
: | |
Db 8 WLLALATWHTSAYGPDROAKKGDIIIGLGEPIHF--GVAARD--ODLKSRESVVCIR 62
: | |
QY 69 YNL-----MOAMRAVEEINNDSSLPLGVLGYEIVDVC-YISNNQPVLYPLA-HEMDL 121
: | |
: | |
Db 63 YNRFGRWLOAMTFAIEINSSPALPLNLGLTIFTCNTVSALATTSFAQONKIDS 122
: | |
QY 122 LPIDQ--DYSNYISRYVAVIGPDPNSESMTVANFLSLFLPQITYSASIDELRDKVFP 179
: | |
: | |
Db 123 LNDEFNCSEHPISTAVVAGATSGSTAVANLGLFIYIPQVYASSRLSKNQFKS 182
: | |
QY 180 LKRTPSADHVEAMVOMLMEFRNMTIIVYSSPTVGRDNGQLGGERVARDICIAQET 239
: | |
: | |
Db 183 FLRTIPNDEHQAATMADIIIEFRNNWVGTAADDYGRPGIIEKREAEERDIDICIFSEL 242
: | |
QY 240 LPITQPNQNTSEEROLATYIVDKLOOSTARVVVSPDLTVLHFEVLRONFTGAVMI 299
: | |
: | |
Db 243 I-----SQSDEEITQVAVEYI---QNSTAKVIYVSSGDLLEPLINEIYRNITGKIWL 294
: | |
QY 300 ASESMAIDPYLHNLTEGLHGLTIGITIQSVPIPGFSEF-----REMGPO 344
: | |
: | |
Db 295 ASEAMASSSLIAMPQYHVAVGGTIGFALKAGQJPGFREFLKVKVPRKSVANGFAKEWEE 354
: | |
QY 345 A-----GPPPLSTQSTYCNQCDNCLNATISFTIILRLSE-----382
: | |
: | |
Db 355 TFNCHLOEGAKGPLVDTFLRGH--EESGDRFNSSTAFRPL--CTGDENISSVEPYID 410
: | |
QY 383 -----RYVYVYSAVAVAHLSLGC-----DKSTCRVYVPMOLLEETWKNFT 430
: | |
: | |
Db 411 YTHLRISYNYLAVYSTAHALODIYCTLPGRGLFTNGSCADIKKAVAMOVKLRLHRLNFT 470
: | |
QY 431 -LDHQIFPDPOGDVALHLEIYOMOMDRSONP--FOSVASYPL-----OROLKNIODIS 482
: | |
: | |
Db 471 NNGGEQVTFPECGDLVGNSTIIMHLSPEBGSIFYKRV--GYVNVYAKKGRFLINEEKL 529
: | |
QY 483 WHYVNNITPMSCKRCQSQOKKPV-GIHVCECEIDCLPGFLNTHTEDEYECQACPN 541

```

```

Db      530  MSGSRREVFPSCNRDCLAGTRKGIIGEPCCCECEVCECPDGETSDBT--DASACNCPDD 5888
      542  FWSYQSTESCCKROLVLEHWEHAPTIAVALLAALGELSTAILIYEMRHQPIVRSAGG 6011
Qy      589  FWSMNEHTSCIAXIEFLSWTEPBGIALTLFAVIGLTFALFVGFKEFKENTPIVAKATNR 6488
Db      602  PMCFMLTLILVAAMVVPVYVGPBKVSTCLCRALPFLCTICISCIAVSPQYCAF-- 6599
Qy      649  ELSTLLFSLTLCSSSLFFIGEPQDWTCCRLOPAFISVLCTISCLIVTNRNLLFEA 7080
Db      660  KMAERFPRAYSYVWRIOGOPYSMAFIVLKMVYIVIGMLARPOSHPTDDDDPKRTIVSC 7190
Qy      709  KIPSPFHKK---WGLNMQFLVFLCIPMQIYICVIMLYLAPBSYSTYNOLEDEBITITTC 7555
Db      720  NPNTRNLSLENTSIDLLISVGVGSFAYMGKELPTYNNEAKFTILSMTEYFTSSVLCTEM 7799
Qy      766  HEGSILMALGFIQYTCILAAICFEFAKRSKRLPENENAEKFIETSMILFIPIWIS---FI 8222
Db      780  SAYSGVLTVIIDLVTYVNLNLATISLG---YFGKCYMILFYPERNRT 8222
Db      823  PAIYSTYGEKVS-AVEVATIAAASEGLLACIFENKKIITILFEKPSRNT 8688

```

RESULT 6
US-08-484-565-7
Sequence 7, Application US/08484565
Patent No. 5763569
GENERAL INFORMATION:
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,565
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451

```

? FILING DATE: 23 August, 1991
? ATTORNEY/AGENT INFORMATION:
? NAME: Heber, Sheldon O
? REGISTRATION NUMBER: 38,179
? REFERENCE/POCKET NUMBER: 213/006
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (213) 489-1600
? TELEFAX: (213) 955-0440
? TELEX: 67-3510
? INFORMATION FOR SEQ ID NO: 7:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1078 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: protein
?
?
? US-08-484-565-7

```

QY 720 NPNYRNSLNTSLDLLSVGFSFAYMGKELPTNYNEAKFTLMTFTSVSLCTFM 779
 Db 766 HEGSIMALGFLIGYTCILAAICFEFFAKSRKLPENFNBAKFTSMILFTVWIS---FI 822
 QY 780 SAYSGVLTIVDLYVNLNLAISLG---YFGKCYMILFYPRNT 822
 Db 823 PAVASTYGRKFS-AVEVIAIILAAISFGILACIFKFKIYIILFKPSRNT 868

RESULT 7

US-08-480-751-7
 ; Sequence 7, Application US/08480751
 ; Patent No. 583684
 ; GENERAL INFORMATION:
 ; APPLICANT: Edward F. Nemeth
 ; APPLICANT: Edward M. Brown
 ; APPLICANT: Steven C. Hebert
 ; APPLICANT: Forrest H. Fuller
 ; APPLICANT: James E. Garrett, Jr.
 ; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lyon & Lyon
 ; STREET: First Interstate World Center
 ; STREET: Suite 4700
 ; STREET: 633 West Fifth Street
 ; CITY: Los Angeles
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 90071
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: FASTSEQ
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/480,751
 ; FILING DATE: 7 June, 1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; PRIOR APPLICATION DATA: including application
 ; PRIOR APPLICATION DATA: described below: 9
 ; APPLICATION NUMBER: 08/353,784
 ; FILING DATE: 9 December, 1994
 ; APPLICATION NUMBER: PCY/US/94/12117
 ; FILING DATE: 21 October, 1994
 ; APPLICATION NUMBER: U.S. 08/292,827
 ; FILING DATE: 23 August, 1994
 ; APPLICATION NUMBER: U.S. 08/141,248
 ; FILING DATE: 22 October, 1993
 ; APPLICATION NUMBER: U.S. 08/009,389
 ; FILING DATE: 23 February, 1993
 ; APPLICATION NUMBER: U.S. 08/017,127
 ; FILING DATE: 12 February, 1993
 ; APPLICATION NUMBER: U.S. 07/934,161
 ; FILING DATE: 21 August, 1992
 ; APPLICATION NUMBER: U.S. 07/834,044
 ; FILING DATE: 11 February, 1992
 ; APPLICATION NUMBER: U.S. 07/749,451
 ; FILING DATE: 23 August, 1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hebert, Sheldon O.
 ; REGISTRATION NUMBER: 38,179
 ; REFERENCE/DOCKET NUMBER: 213/004
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (213) 489-1600
 ; TELEFAX: (213) 955-0440
 ; TELEX: 67-3510
 ; INFORMATION FOR SEQ ID NO: 7:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1078 amino acids
 ; TYPE: amino acid

TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-480-751-7

Query Match 24.6%; Score 1094; DB 2; Length 1078;
 Best Local Similarity 31.5%; Pred. No. 6,7e-95;
 Matches 279; Conservative 160; Mismatches 342; Indels 106; Gaps 26;

QY 16 WVLAEPAENSEFLP-----GDYLLGLFSLHANKGIYHNLFLQVPCKEVEYVIG 68
 Db 8 WVLALTWHITSAYGPDOAROKKDDIILGLFPPIHF---GVAARD--QDLKRSPEVDCIR 62
 QY 69 YNL-----MQAMRAVEEINNDSSLPGLVEIYDVC-YISNNQVPVLYFLA-HEDNL 121
 Db 63 YNFRGFRMLQAMIFAIKEINSSPALPLNLLGYRIFPTCVTSKALFATLSFAVNKIDS 122
 QY 122 LPIDF--DYSNYISRVAVVGPDSSEVMYVANTSLFLPQITYSAISELBDKVFPA 179
 Db 123 LNLDEFCNSEHLPSTLAIVAGGSGVSTAVANLLGLFYIPQVYASSRLLSKNQFS 182
 QY 180 LRLTPSADHHEAMVOLMLFRNMWIIIVSSPTYGRDNGQLGEHVARDDICIAFOET 239
 Db 183 FLRTIPDEHOATMADIIEFRNMWGTIADDDYRPGIEKREAEERDIDISEL 242
 QY 240 LPTLQPNQNTSEERQLVTVDKLQOSTARVVVSPDLTLHFNEVLRFNFTGAVMI 299
 Db 243 I-----SQYSEDEHIOHVEVI---QNSTAKVIIVFSSGPDLEPLIKEIYRRNTGKIWL 294
 QY 300 ASESMALDPVILNTELGHTFLGITIQSVPIGSESEF-----REMGPO 344
 Db 295 ASEMASSSLIAMPQYFHVVGTTGFALKAQOIPGFEFELKVKYPRKSVHNGFAKEWEE 354
 QY 345 A-----GPPILSRYSQSYTCNOECNCLNATISFTLILSGE----- 382
 Db 355 TFNCHLOEAGKPLPVTFILRGH--EESGDRFSNSIAFAPL--CTGDENISSVEPYID 410
 QY 383 ---RVVYSYSAVYAAVAHLSLGC-----DKSTCKRVVYFWLLIEIMKYNFT 430
 Db 411 YTHLRISYNYLAVYSTAHALODIYTCPLPGRLFTNSCADIKKEVAMQVYLKHLRLNFT 470
 QY 431 -LIDHOLFPPQGVNALHLEIYQWMDRSONP--FQSVASYPL-----OROLKNIDIS 482
 Db 471 NNMGEOVTFDECGDLVGNYSITNNHLSPEDEGSIVFEKV--GYVNYAKKGERLPTNEKIL 529
 QY 483 WHTVNTTIPMSMCKRSQSQKKRPV-GIHYCEPCIDCLPGFLNTHTEDEYECQACPPN 541
 Db 530 WSGFSRREVPVNSCRDCLAGRKGIIEGPTCECEVCEPDGEYSDET-DASACNKPDD 588
 QY 542 EWSYQSETSCPKROLVLEHNEAFTIYVALLAALGFLSTIALIVTFRRHQTPIVRSAG 601
 Db 589 FWSNENHTSCIAKIEIEFLSWTEPGIALTFEAVIGIFLTFVGLVFIFKFRNTPIVKATNR 648
 QY 602 PNCPLMTLLLVAMVVPVAVGPKVSTCROALPLFCITGICSLAVRSFOVCAF-- 659
 Db 649 ELSTLLFLSLCCPSSSLFTIGEPDWTCLRKQPAFISIVYLCSCLVATNRLLVFEA 708
 QY 660 KMASRPRAYSYWRYOGPVYMAFITVLMKVIYVIGMLARPOSHPTDDDPKTIIVSC 719
 Db 709 KIPISFHRK---WNGMLQGLVFLFCFMQIVICVIMLYIAPSPSYNNOLEDEBILFITC 765
 QY 720 NPNYRNSLNTSLDLLSVGFSFAYMGKELPTNYNEAKFTLMTFTSVSLCTFM 779
 Db 766 HEGSIMALGFLIGYTCILAAICFEFFAKSRKLPENFNBAKFTSMILFTVWIS---FI 822
 QY 780 SAYSGVLTIVDLYVNLNLAISLG---YFGKCYMILFYPRNT 822
 Db 823 PAVASTYGRKFS-AVEVIAIILAAISFGILACIFKFKIYIILFKPSRNT 868

RESULT 8

US-08-943-986-7
 ; Sequence 7, Application US/08943986
 ; Patent No. 5962314

```

GENERAL INFORMATION:
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,986
FILING DATE: 03-OCT-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/484,565
FILING DATE: 7-June-1995
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Hebert, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1078 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-943-986-7

Query Match 24.6%; Score 1094; DB 2; Length 1078;
Best Local Similarity 31.5%; Pred. No. 6.7e-95;
Matches 279; Conservative 160; Mismatches 342; Indels 106; Gaps 26;

```

```

122 LPIQ--DYSNYSRVAVIGPDNSESVTANFLSLFLLPQITYSATSDLRDVRIPA 179
123 LNLDFCNSEIIPSTIAVVGATGSGVSTAVANLGLFIPVYASSRLLSNKQFRS 182
140 LIRTPSADHVEAVQMLHFRMWTIVLVSSDYGRDNGOLLGERVARDICIAFOET 239
143 FLRTIPNDHQATAMADILEYFRMMWVGIIAADDDYGRGRIKFRFEAEERDIDFSEL 242
240 LPTLPQNMNISEQRQLTYDKLQOSRVRVYFSPDLTIYHFNELRQNFQAWI 299
243 I-----SOYSDEEIOHVEVI---QNSTAKYIVFSSGPDLEPLIKETVRNITGKWL 294
300 ASESNAIPVLNHLNTELGITGTTQSYVIPFSEF-----REMGPO 344
295 ASEANASSLIMPOYFHVVGSTIFALKAGIIPGFRFLKRVHRRKSVHNFAPAEFNEE 354
345 A-----GPPLSRTSQSYTCNQBCDCLNATLSFNTLRSGE----- 382
345 TPNCHLQEGAKGPLVDFELRGH--EESGDRFSNSFARPL--CTGDENISSVEPYID 410
363 ----KIVSYSAVAVAHALSLGC-----DKSICTRVYYPQMLEIKVNF 430
411 YTHLRISYNNVLAVYSIAHALODIYTCLRGRLFTNGSCADIKYEAQVILKHLNFT 470
431 -LLDHOIFPDPOGDVALHLEIVQMOMDRSQNP--FOSVASYPL-----OROLKNIODIS 482
471 NMGEOVTFDECGDLVGNSTIINMHLSPEDGSIVREV--GYNNVAKKGERLFINBEKIL 529
483 WHTVNTIIPMSCKRCSGQAKRPV-GIHWCCFECIDCLPSTPLNHEDEYECQACPNN 541
530 WSGFREVDFNSCRDCLAGTRKGIIEGPECPFCFECVDEPDEYDET--DASACNKKCPD 588
542 EWSYSEISCFKROLVLEHMAPITIAVALALGFLSLALIVFMRHFOPIYRSAG 601
589 FMSNNHTSCIAKEIEFLSWTEPGIALTLEFIPLAFGLVGIKFRNPIYKATNR 648
602 PRCFLMLTLIAVAVVPPVYGPVPSYICQALFPLCTFCICIAVRSQIACAF-- 659
649 ELSTYLLSLCCFSSSEFIEPQDWTCLRQPAFISFVICISICILKTRVILYPERA 708
709 KIPTSFHRK---WMLNLQFLFLVFLCTEQWIVCIWLYTAPPSYRNQLEDEILFIC 765
720 NNTNNSLLENTSLDLSVGFSPAYMKELEPTYNNAKFTLSMTFTTSYSVLCTFM 779
766 HEGSIMALGFLIGYTCLLAALICFFFAFRSKLPEFNENKRTFTFSMLIFVIWIS--FI 822
780 SAYSGVLVYIDLTVYTNLILASISG---YFPGKCYMILFPERNT 822
823 PAYASTYCKFVS-AVEVAIIILASFGLLACIFFNKIIITLFRPSNT. 868

RESULT 9
US-08-353-784-7
Sequence 7, Application US/08353784
Patent No. 6011068
GENERAL INFORMATION:
APPLICANT: Edward F. Nemeth, Edward M.
APPLICANT: Brown, Steven C. Hebert,
APPLICANT: Bradford C. Van Wageningen, Manuel
APPLICANT: F. Balandrin, Forrest H. Fuller,
APPLICANT: Eric G. Delmar, and Scott T. Moe
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles

```

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000

```

```

: FILING DATE: 9 December, 1994
: APPLICATION NUMBER: PCT/US/94/12117
: FILING DATE: 21 October, 1994
: APPLICATION NUMBER: U.S. 08/292,827
: FILING DATE: 23 August, 1994
: APPLICATION NUMBER: U.S. 08/141,248
: FILING DATE: 22 October, 1993
: APPLICATION NUMBER: U.S. 08/009,389
: FILING DATE: 23 February, 1993
: APPLICATION NUMBER: U.S. 08/017,127
: FILING DATE: 12 February, 1993
: APPLICATION NUMBER: U.S. 07/934,161
: FILING DATE: 21 August, 1992
: APPLICATION NUMBER: U.S. 07/834,044
: FILING DATE: 11 February, 1992
: APPLICATION NUMBER: U.S. 07/749,451
: FILING DATE: 23 August, 1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Douglas C. Murdock
: REGISTRATION NUMBER: 37,549
: REFERENCE/DOCKET NUMBER: 213/007
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELEX: 67-3510
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1078 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-484-7198-7

```

```

Query Match      24.6%; Score 1094; DB 3; Length 1078;
Best Local Similarity 31.5%; Pred. No. 6.7e-95;
Matches 279; Conservative 160; Mismatches 342; Indels 106; Gaps 26;

```

```

QY 16 WTLAPAEKNSDEYLP-----GDYLLGLGFLSLHNMKGIVLNLFLQYPMCKEYKVG 68
DB 8 WLLALTLHTTSAYGPDRQAKKGDIILGLFPIHF---GVAANKD--QDIKSPRESVECHR 62
QY 69 YNL-----MQAMRFAVEEELINDSSLPGVLLGYEIVDYC-YISNNVQPLYELA-HEENL 121
DB 63 YNRFGRFLQAMIFAEELINSPLALNLTGLRIFDTCNTVSKALEATLSVAQNKIDS 122
QY 122 LPIDQ--DYSNTISKRVAVIGPDNSESVATVANFLSLPDIPTTSAISDELDRKVRPA 179
DB 123 INLDEFCNCSEHIFSTIAVAGTSGSVTAANLGLFYIPQVSYASSRLLSNKQFES 182
QY 180 LRTTPSADHVEAMVQLLHFRMWIIYLVSSDTYGRDNGOLLGERVARRDICIAFORT 239
DB 183 FLRTIPNDHQATAMADITETFRMNVGIIAADDDYGRGIEKFRFEAEERICIDFSEL 242
QY 240 LPTLPQNMQNTSEERQRLTVTKLQOSTARVVVFPDLTLTHFENEVLKRONFTGAWVI 299
DB 243 I-----SOYSDDEEIQHVEVI---QNSTAKVIYVSSGPDLEPLIKEIVRRNIGKIML 294
QY 300 ASESNAIDPVYLNHLDELGLGFTLTIOSVPIPGFSEF-----KEMGRQ 344
DB 295 ASEAWASSSLIMPOYFHVVGTTIGFPAKAGQIPGRFELKRVHPRKSVHNGFAEFWE 354
QY 345 A-----GPPLSRTSQSYTCNQECNCLNATLSFNTILRLSGE----- 382
DB 355 TPNCHLOEAKGRLPVDPLRGRH--ESGDRFSNSTARPL--CTGDNISSVETPYID 410
QY 383 ----RVYSVYSAYVAVAHALSLIGC-----DKSTCKRVVYPMQLLEIMKVNFT 430
DB 411 YTHRLISVYVYLAVYSIAHALODITYCLPGRGLFTNGSCADIKKYEAMQVHLKHLNFT 470
QY 431 -LLDHOIFEDPDQGVALHLEIYQWQMDRSONP--FQSVASYRPL-----QRLKNIQDLS 482
DB 471 NNMGEQVYTFDECGDLVGNSTITNMHLSPPDGSIYFREV--GYINVYAKKGERLPINEKITL 529

```

```

QY 483 WHTVNTTIPMSCKRCOSGOKKKPV-GIHVCCFRCIDLPSTFLNHTDEYEQACAPRN 541
DB 530 WSGFREFRPEVFNCSRDCLAGTRKGIIIEGPTCCFCEVCEPDDEYSDET--DASACKNCPDD 588
QY 542 EMSYQSESCFKROLVFLFEMHEAPTLAVALLAALGFLSTLALIVFWRHFOPIYRSAGG 601
DB 589 FSNENHISCIKAEIEFLSWTEPRGIALTLFPAVLGIFLPAVLAGYIKRNRNPIYKATNR 648
QY 602 PRCFLMLTLLVAVYVVPYVGPVKVSTCLCQALPELCTTICISCIAYRSFOYCAF-- 659
DB 649 ELSTYLLFELLCFSSSEFFIGEPQDMTCRLQPAFGISFVLCISCIITVKTNRVLLVFEBA 708
QY 660 KNASRPRATSYWVRVYQGGYVSMATFVTKMVIYIGMLARQSHRPDPDDPKTIYVSC 719
DB 709 KIPTSFHRK---WGLNINQFLVPLCTEWOIVCIWLTATAPSSYRNOEDELDELITFTC 765
QY 720 NPNYRSLFNTSLDLLSVGFSFAYMGEKLPJYVNEAKFTLLSMTEFFTSVSLCTPM 779
DB 766 HEGSLMALGFLGYTCILAAICFFPAFKRKLPENFNEAKFTTFSMLIFFIYWIS---FI 822
QY 780 SAYSGVLYTVIDLTVNLALISLG---YGPCCYMLTFYPERNT 822
DB 823 PAYASTYGRFVS-AVEYAIILAAISFGLACIFPNKIYIILLFRPSRWT 868

```

RESULT 11

```

US-08-484-159-7
; Sequence 7, Application US/08484159
; Patent No. 6313146

```

```

; GENERAL INFORMATION:
; APPLICANT: Bradford C. Van Wageningen
; APPLICANT: Manuel F. Balandrin
; APPLICANT: Eric G. Del Mar
; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
; TITLE OF INVENTION: MOLECULES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: First Interstate World Center
; STREET: Suite 4700
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA

```

```

; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FASTSEQ

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,159
; FILING DATE: 7 June, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below: 9

```

```

; APPLICATION NUMBER: 08/353,784
; FILING DATE: 9 December, 1994
; APPLICATION NUMBER: PCT/US/94/12117
; FILING DATE: 21 October, 1994
; APPLICATION NUMBER: U.S. 08/292,827
; FILING DATE: 23 August, 1994
; APPLICATION NUMBER: U.S. 08/141,248
; FILING DATE: 22 October, 1993
; APPLICATION NUMBER: U.S. 08/009,389
; FILING DATE: 23 February, 1993
; APPLICATION NUMBER: U.S. 08/017,127
; FILING DATE: 12 February, 1993
; APPLICATION NUMBER: U.S. 07/934,161
; FILING DATE: 21 August, 1992

```



```

: APPLICATION NUMBER: U.S. 07/834,044
: FILING DATE: 11 February, 1992
: APPLICATION NUMBER: U.S. 07/749,451
: FILING DATE: 23 August, 1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Hebert, Sheldon O.
: REGISTRATION NUMBER: 38,179
: REFERENCE/DOCKET NUMBER: 214/101
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELEX: 67-3510
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1078 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-484-159-7

```

```

Query Match      24.6%; Score 1094; DB 4; Length 1078;
Best Local Similarity 31.5%; Pred. No. 6,7e-95;
Matches 279; Conservative 160; Mismatches 342; Indels 106; Gaps 26;

```

```

QY 16 WTAEPAENDSYLP-----GDYLGGLSLHANKGIYHLNFIQYPMCKEYEVKYG 68
D 8 WLLATWHTSAYGPDRQAKKGDIIIGLFPHF---GVAAKD--DOLKSRPSEVCEIR 62
QY 69 YNL-----MQMRFAVEINNDSSLPGVLLGEIVDYC-YISNNQPVLYFLA-HEENL 121
D 63 YNFGFHWLQAMIFALIEINSSPALLPLTLGYRIPTCNVSKALEATLSFVANKIDS 122
QY 122 LPIDQ--DYSNYISRYVAIVIGPDNSESVMYANFLSLFLPQIYSAISDELPRKVRPA 179
D 123 LNDEPCNCSHITSTAVAGATSGSVTAVANLGLFYIPOVYASSRLSKNKQPKS 182
QY 180 LRTTPSADHVEAMVOLMHERNNWITIVVSDTYGDNQQLGERVARDICIAFOET 239
D 183 FLRTIPNDEHQTAMADIEYFRMNVGTIAADDYGRPIEKFREAEEDICIDFSEL 242
QY 240 LPTQPNOMMTSEERORLYTVDKLOOSTARVYVVSFDPDLLHFNFNVLQNTGAWI 299
D 243 I-----SOYSDDEEIOHVEYI--QNSTAKVIVFSSGPDLEPLIKEIVRNITGKWL 294
QY 300 ASESMAIDPVLANLTELGLHGTFLGITIOSVPIDGSEF-----REMGPO 344
D 295 ASEMASSSLAMPQYFHVVGITIGPALKAGQIRGFRFLKVKPRKSVHNGFAKEFWE 354
QY 345 A-----GPPPLSRTSOSTYCNOCNCLNATLSFNTILRLSGE----- 382
D 355 TFNCHLOEGAKGPLVDFTFLRGH--EESGDRFSNNSSTAFLRL--CTGDENISSVETPYID 410
QY 383 -----RVVYSYSAVAYVAHLSLGC-----DKSTCKRKYVPMQLLEETAKVET 430
D 411 YTHRISTNYLAVYSTAHALQDITICLPGRLFTNSCADIKKVEAMQVILKHLRHNET 470
QY 431 -LIDHQIFPDPOGVALAHLLEIVQMDRSONP--FQSYASYPL-----OHLNKIDIS 482
D 471 NNMEQYTFPDCGLDVGNYSIINMHLSPEDOSIYFKEY-GYNNYAKKGELFTINEKITL 529
QY 483 WHYNNNTIPMSCKSKROSGOKKRPV-GIHVCFECIDCLGTFELNHTDEYECQACPN 541
D 530 WSGSRREVPFSCSRDLGTRKGIIEGPTCFCEYECPEGEYSDET-DASACNKPDD 588
QY 542 EMSVOSTSCRKROLVLEHTEAFTIYVALLAAGFLSTAILVIFMRHPTPIYRSAGG 601
D 589 FWSMENTSCIAKEIEFLSWTEPGIALTFEAVIGIFLTAFLVGLFKFRNTPIYKATNR 648
QY 602 PMCIAMITLLIYAVVYVYGPVKSTCLCRQALFPLCFITICISCAVRSFOVCAF-- 659
D 649 ELSTVLLFLSLCCSSSLFTIGEPQDMTCRLRQPAFGISYFLCLISCLIVTKNRLLAFEA 708
QY 660 KMASRFPRAVSYYWYOGPYVSMFITVLKMYIVVIGMLARPOSHPRTPDDPKITIVSC 719

```

```

D 709 KIPTSFARK---WMLNLQFLVFLCTFMQIVICVIMLYTAPPSRYNOELEDEIRITC 765
QY 720 NPNRNSLNFETSIDLLSVYGEFSAFMKGLPNNYNEAFITLSMFFITSSISCTFM 779
D 766 HEGSLMGLIGYTCILAAICFFFAKSRRLPNNFENAFITSSMLFFIWTIS--FI 822
QY 780 SAYSGVLVTYIDLIVYVNLNLAISLG----YFGRCXMIFFEPRT 822
D 823 PAVASTYGRKVS-AVEYIALIILASFGLLACIFFKIITILFKPSRNT 868

```

RESULT 12

```

US-09-162-021B-2
: Sequence 2, Application US/09162021B
: Patent No. 6337391

```

```

: GENERAL INFORMATION:
: APPLICANT: H. William Harris
: APPLICANT: Edward M. Brown
: APPLICANT: Steven C. Hebert
: TITLE OF INVENTION: Polycation-Sensing Receptor in Aquatic
: TITLE OF INVENTION: Species and Methods of Use Thereof
: FILE REFERENCE: 2856.1001-007
: CURRENT APPLICATION NUMBER: US/09/162,021B
: PRIOR FILING DATE: 1998-09-28
: PRIOR APPLICATION NUMBER: PCT/US97/05031
: PRIOR FILING DATE: 1997-03-27
: PRIOR APPLICATION NUMBER: 08/622,738
: PRIOR FILING DATE: 1996-03-27
: NUMBER OF SEQ ID NOS: 19
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 2
: LENGTH: 1027
: TYPE: PRT
: ORGANSIM: squallus acanthias
: US-09-162-021B-2

```

```

Query Match      24.6%; Score 1092.5; DB 4; Length 1027;
Best Local Similarity 31.0%; Pred. No. 8.6e-95;
Matches 277; Conservative 162; Mismatches 360; Indels 95; Gaps 25;

```

```

QY 5 AKTICSEFLIMVTAEPNDSFYLP-----GDYLGGLSLHANK--KGIYHLNFIQ 55
D 2 AQLHCOLFLFGLFTLQ--SYNVSGYPNORAKKGDIIIGLFPHFVGAKKDDOLKSRPE 60
QY 56 VPMCKEYEVKYGYNLQAMFAVEEINNDSSLPGVLLGEIVDYC-YISNNQPVLYE 114
D 61 ATKCIIRYNR--GRWLOAMIFALIEINSSKTPLENTLIGRTIDTCNTVSKALEATLSF 118
QY 115 LA-HEENLPIQE--DYSNYISRYVAIVIGPDNSESVMYANFLSLFLPQIYSAISDEL 171
D 119 VAQNKIDISLNDEPCNCSDHIPSTIYAVGATGSGISYAVANLGLFYIPOVYASSRL 178
QY 172 RDKVFPALLRTTYSADHVEAMVOLMHERNNWITIVVSDTYGDNQQLGERVARD 231
D 179 SNKNEYKAFLETPINDCOQATMAEIEHFQMNWVGTIAADDYGRGIDKFRREAVKRD 238
QY 232 ICIAFOETLQPNOMMTSEERORLYTVDKLOOSTARVYVVSFDPDLLHFNFNVLQ 291
D 239 ICIDFSEMT-----SOYTT--QKQLEFIADVIONSSAKVIVFSSGPDLEPLIQTIVR 290
QY 292 NFGTAVWIASMAIDVLANLTELGLHGTFLGITIOSVPIDGSEF----- 340
D 291 NIDIRIWLASPMASSSLIAPKPEFHVVGITIGFALAGRIPGNKRKLKVEHPSRSSDG 350
QY 341 -----W-----GPOA-GPPPLSRTSOSTYCNOCNCLNAT 370
D 351 FVKEFEWETPNCFYETKLTITOLKNSKVPSHGPAQDGSRAKNSRRRLAHPCTGSEENIT 410
QY 371 LSFPTILRLGERVYSVYSAVYVAHLSLIGCDST-----CTK--RVYVPMQLE 422
D 411 SVETPYLDYTHLRISYNNVAVYSIAHALQDISHCKPGTGFIFANGSCADIKKVEAMQVLN 470

```


Db 642 IYKATNELSYLLFLSLCFSSSLFIEGPODWTCLROPAGISFVLCISCLIVYTNR 701
 QY 655 IYCAF--KMASRPRASVWRYGQPVYSAFTVLMVIVIGMLARPOSHPTDDP 712
 Db 702 VLVFEAKIPTSFHRK--WMLGLOFLVFLCTFMOLICTIWLPAAPSSYNHLED 758
 QY 713 KITVSCNPVNRNSLFTNTSL--DLLSVGFSFAYMGKELPTNYEAKETILSMTEYF 769
 Db 759 EIIPTTC--HEGSLMALGSLIGYTCLLAICFFAFKSRKLPENFNAKFTSMILIF 815
 QY 770 TTSVSLCTFMSAYSGVLVTIVDLVTVNLTAISLG---YFGKCYMILFYPERNT 822
 Db 816 IWTIS--FIPAVASTYGRKFS-AVEYIALAASFGILACIFFENKYYIILFKPSRNT 868

 RESULT 14
 US-08-484-565-8
 Sequence 8, Application US/08484565
 Patent No. 5763569
 GENERAL INFORMATION:
 APPLICANT: Edward M. Brown
 APPLICANT: Steven C. Hebert
 APPLICANT: James E. Garrett, Jr.
 TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 TITLE OF INVENTION: MOLECULES
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: First Interstate World Center
 STREET: Suite 4700
 STREET: 633 West Fifth Street
 CITY: Los Angeles
 STATE: California
 COUNTRY: USA
 ZIP: 90071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: FASTSEQ
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/484,565
 FILING DATE: 7 June, 1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA: including application
 PRIOR APPLICATION DATA: described below: 9
 APPLICATION NUMBER: 08/353,784
 FILING DATE: 9 December, 1994
 APPLICATION NUMBER: PCT/US/94/12117
 FILING DATE: 21 October, 1994
 APPLICATION NUMBER: U.S. 08/292,827
 FILING DATE: 23 August, 1994
 APPLICATION NUMBER: U.S. 08/141,248
 FILING DATE: 22 October, 1993
 APPLICATION NUMBER: U.S. 08/009,389
 FILING DATE: 23 February, 1993
 APPLICATION NUMBER: U.S. 08/017,127
 FILING DATE: 12 February, 1993
 APPLICATION NUMBER: U.S. 07/934,161
 FILING DATE: 21 August, 1992
 APPLICATION NUMBER: U.S. 07/834,044
 FILING DATE: 11 February, 1992
 APPLICATION NUMBER: U.S. 07/749,451
 FILING DATE: 23 August, 1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Hebert, Sheldon O.
 REGISTRATION NUMBER: 38,179
 REFERENCE/DOCKET NUMBER: 213/006
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510

; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1079 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-484-565-8

 Query Match 24.5%; Score 1088.5; DB 1; Length 1079;
 Best Local Similarity 31.7%; Pred. No. 2.2e-94;
 Matches 284; Conservative 164; Mismatches 328; Indels 121; Gaps 31;

 Db 9 ALALAN-----HSSAYGPDQRAKKGDIIIGGLPIH---GVAKD--QDLKSRPE 56
 QY 63 EVKIVGYNL-----MQAMRAVEEINDDSLPGVLLGLEYVDVC-YISNVQPVLYFLA 116
 Db 57 SVECIKRYNFGFWLQAMIRAEIENSPPSLPMTLGYRIFDCTNVSALNTLSFVA 116
 QY 117 -HEDNLPDIOE--DYSNYSRVAVAVIGPDNSESVAFVAFSLFLPQITYAISDELARD 173
 Db 117 QNKIDSLNDEFCSCSHISTIVAGATGSGVSTANALGLFYIYQVYASSRLSN 176
 QY 174 KVFPELLRTTPSADHVEAVQMLHFRRNMTIVLVSSPTGRCNQLGERRYARDIC 233
 Db 177 KQYKSEFLRTIPNDHQATMAADIEFFRRNWVGTIAADDYGRGIEKFEKEEEDIC 236
 QY 234 IAPQETLPTLOPNOMNTSEORLVTVDKLOOSTARVVVFPDLDLHFNFVTLRQNF 293
 Db 237 IDESELI-----SOYSDEELQOVVEVI--QNSTAVIVFSSGPDLEPLIKETVRNI 288
 QY 294 TGAVWIASGWAIDPVLANTELGH-LGTELGITYQSVPIGFSF-----338
 Db 289 TGRWILASEAMASSSL-ANPEYHVAVGRTIGFGLKAGQIPGFEFFQXVHPRKSVANGF 347
 QY 339 -REWGPOA-----GPPUSRTQSSTYCNQECNCLNATLSFTILRSGE-----382
 Db 348 AKERFEETFNCHLOEGAKGLPVDTFVRSH--EEGGRNLLNSVAFRPL--CTGDEINNS 403
 QY 383 -----RVVYSVYVAVAHALHSLGC-----DKSCTKRVYVMOLLE 423
 Db 404 VEPYMDYEHRLRISTYNYLANVYSTAHALODIYTCLPRGLFTNSCADIKVEAMOVLYKH 463
 QY 424 IWKNFT-LDHOJFPPQSDVALHLEIVOMWDRSONP--FQSVASYPL-----OROL 475
 Db 464 LRHLNFTNMNGEYVTFDECGDLVGNYSIIMWHLSPEDGSIYKEV--GYVNVYAKKGERLF 522
 QY 476 KNIDDSMHVYNNITIPMSCKSRQSQOKKPV-GIHYCCFECIDCLPGFLNTEDEYE 534
 Db 523 INEKILMSGFSREVPSNCRDQACTRGIIIEGPTCCFCEVCECDGEYSGET--DASA 581
 QY 535 QOCPNNMWSYQSESCFRLQVLEFWHEAPITVALAALGFLSTALIVYEFRRHPTP 594
 Db 582 CDCRPPDFMSNENTSCIAEIEFLANTPERGIALTFPAVLGIVLAFVGLFKRNTP 641
 QY 595 IYRSGGPMCLMTLLLVAVMVVYVGPVKSTCLCROALPELCTICIAIVBSFQ 654
 Db 642 IYKATNELSYLLFLSLCFSSSLFIEGPODWTCLROPAGISFVLCISCLIVYTNR 701
 QY 655 IYCAF--KMASRPRASVWRYGQPVYSAFTVLMVIVIGMLARPOSHPTDDP 712
 Db 702 VLVFEAKIPTSFHRK--WMLGLOFLVFLCTFMOLICTIWLPAAPSSYNHLED 758
 QY 713 KITVSCNPVNRNSLFTNTSL--DLLSVGFSFAYMGKELPTNYEAKETILSMTEYF 769
 Db 759 EIIPTTC--HEGSLMALGSLIGYTCLLAICFFAFKSRKLPENFNAKFTSMILIF 815
 QY 770 TTSVSLCTFMSAYSGVLVTIVDLVTVNLTAISLG---YFGKCYMILFYPERNT 822
 Db 816 IWTIS--FIPAVASTYGRKFS-AVEYIALAASFGILACIFFENKYYIILFKPSRNT 868

RESULT 15
 US-08-480-751-8
 : Sequence 8, Application US/08480751
 : Patent No. 5858684
 :
 : GENERAL INFORMATION:
 : APPLICANT: Edward F. Nemeth
 : APPLICANT: Edward M. Brown
 : APPLICANT: Steven C. Hebert
 : APPLICANT: Forrest H. Fuller
 : APPLICANT: James E. Garrett, Jr.
 : TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 : MOLECULES
 : NUMBER OF SEQUENCES: 20
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Lyon & Lyon
 : STREET: First Interstate World Center
 : STREET: Suite 4700
 : STREET: 633 West Fifth Street
 : CITY: Los Angeles
 : STATE: California
 : COUNTRY: USA
 : ZIP: 90071
 :
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: FASTSEQ
 :
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/480,751
 : FILING DATE: 7 June, 1995
 : CLASSIFICATION: 435
 :
 : PRIOR APPLICATION DATA:
 : PRIOR APPLICATION DATA: including application
 : PRIOR APPLICATION DATA: described below: 9
 : APPLICATION NUMBER: 08/353,784
 : FILING DATE: 9 December, 1994
 : APPLICATION NUMBER: PCT/US/94/12117
 : FILING DATE: 21 October, 1994
 : APPLICATION NUMBER: U.S. 08/292,827
 : FILING DATE: 23 August, 1994
 : APPLICATION NUMBER: U.S. 08/141,248
 : FILING DATE: 22 October, 1993
 : APPLICATION NUMBER: U.S. 08/009,389
 : FILING DATE: 23 February, 1993
 : APPLICATION NUMBER: U.S. 08/017,127
 : FILING DATE: 12 February, 1993
 : APPLICATION NUMBER: U.S. 07/934,161
 : FILING DATE: 21 August, 1992
 : APPLICATION NUMBER: U.S. 07/834,044
 : FILING DATE: 11 February, 1992
 : APPLICATION NUMBER: U.S. 07/749,451
 : FILING DATE: 23 August, 1991
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Hebert, Sheldon O.
 : REGISTRATION NUMBER: 38,179
 : REFERENCE/DOCKET NUMBER: 213/004
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (213) 489-1600
 : TELEFAX: (213) 955-0440
 : TELEX: 67-3510
 : INFORMATION FOR SEQ ID NO: 8:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 1079 amino acids
 : TYPE: amino acid
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : US-08-480-751-8

Query Match 24.5%; Score 1088.5; DB 2; Length 1079;
 Best Local Similarity 31.7%; Pred. No. 2,2e-94;
 Matches 284; Conservative 164; Mismatches 328; Indels 121; Gaps 31;
 QY 10 SLFLLWVLAEPANSDYLP-----GDYLLGLGFSIHAANKGIVHLNLFQVPMCKEY 62

```

Db      9 ALNALAW-----HSSAYGPDQRAQRKGDIIIGLGFPIHF---GVAAKD--QDLAKRPE 56
QY      63 EVKVIQYNL-----MQANFAVEEININDSLLPGVLLAGEYIVDVC-YISNNVQPVLYELA 116
Db      57 SVECIKRYNFRGFRWLOAMIFALEEINSSPLLNNTLGRIFDQCNVSKALEATLSFVA 116
QY      117 -HEMDLPIQE--DYSNISRVAIVIGPDNSQSVTVANFLSLFLPQITSAISDELDR 173
Db      117 QNKIDSLNDEFCNCSSEHPISTIAVAGATGSGVSTAVANLGLFTIPQVTSRSLLSN 176
QY      174 KVFEPALRTTPSADHVEAMVQMLAFPMWNIIVLVESDTPYGRDNGQLGERVARRDIC 233
Db      177 KNOYSFLRTIPNDEHQATMAADIEYFRMNVGVIAADDQGRGRIKIFPREAERERIC 236
QY      234 IAFQETLPLQPNQMTSEERORLTVYDKLQOASRARRYVPSDLTIYHFFNEVLRONF 293
Db      237 IDFSLI-----SOYSDEETIQOVVEV---QNSTAKYIVVFSSGPDLEPLIKETVRNI 288
QY      294 TGAVWIAESNAIDPVNLNLFELGH-LGPFGLITTSQVPIPGFSF----- 338
Db      289 TGRITLASEANASSLI-AMPEYFHVVGCTIGFGAKAQIFGFRFLQVHPRKSVHNGF 347
QY      339 -REMGPOA-----GPPPLSRTSOSYTCNQCDCNLTATLSFNTILRSGE----- 382
Db      348 AKEFWEETFNCHLQEGAKGPLVDQFVRSH--EEGNNLNLSSTAFRL--CTGDENINS 403
QY      383 -----RVYVSYSANYAVAAHLSLGC-----DKSTCKRRVYPMQLLBE 423
Db      404 VETPYMDYEHRLISYNYLVAYSIAHALQDIYTCPLGRGLFTNGSCADIKKEVAQVYK 463
QY      424 IMKVNET-LDHOJFEDDQGVALHLEIVOMQMDNSQNP--FQSVASYRPL-----QHQ 475
Db      464 LRLHLEFTNMGEQVTFDDCGDLYGNTSTINHLSPEDGSIVFKEY-GYNNYAKGELF 522
QY      476 KNIDISWHTVNNITPMSCSKRCSGOKKRPV-GIHYCECEIDCLPGLFNLHTEDEYE 534
Db      523 INEEKILMSGFSREVPFNSGRDQAGTRKGIIEGPECCFECVCYCPDGEYSGET-DASA 581
QY      535 COACNNNEMSYOSFSCRRKQRLVFLMEHAPRIAVALLAAGFLSTLAILVIFMHPOTP 594
Db      582 CDKCPDDFWSNENHTSCAKLEFLAWTEPRIALTLFAVIGIFLFAVLGFIIFRMT 641
QY      595 IVRSAGPMCFMLTLILVAVVVPVYGPVSTICRQALFPLCFITICISCIIVRSQ 654
Db      642 IVKATNELSTILLFSLCCFSSSLFTIGEPQDTCRLRQAPFGISFVLCISCIIVKTR 701
QY      655 IVCAF--KMASRFPRAYSYVWRYGQPVYSMAFIVLKVIVYIGMLARPOSHRTDPDP 712
Db      702 VLVVEAKIPTSFHRK---WMGLNQFLVFLCTFMQILICIIMLYTAPSSYRNHELED 758
QY      713 KITIVSCPNTRNSLLENTSL---DLLSVYGFSAYNGKELPTYNNAKFTTLEMTYEF 769
Db      759 EILFTTC---HEGSLMALGSLGYVCLLAICOFFPAFRKRIPEFNNAKFTTSMLETF 815
QY      770 TSSVSLCTFMSAYSGVLTVYIDLTVYVNLALISG---YFGPCYMLFPEPENT 822
Db      816 IYWIS---FIRAYASTYGRFVS-AVEVIAIILASGLACIFPNKYVITILFRPSNNT 868

```

Search completed: May 19, 2003, 09:53:11
 Job time : 18.3799 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:51:25 ; Search time 38.6769 Seconds
(without alignments)
2089.767 Million cell updates/sec

Title: US-09-927-315-9
Perfect score: 4443
Sequence: 1 MGPRKATCSLFFLLWVLAIE.....ERNTPAYENMIGYTRRD 838

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 362588 seqs, 96450795 residues

Total number of hits satisfying chosen parameters: 362588

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_AA:*
1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB pep:*
2: /cgn2_6/ptodata/2/pubppaa/PCT_NEW_PUB pep:*
3: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB pep:*
4: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB pep:*
5: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB pep:*
6: /cgn2_6/ptodata/2/pubppaa/US07_PUBCOMB pep:*
7: /cgn2_6/ptodata/2/pubppaa/PCTUS_PUBCOMB pep:*
8: /cgn2_6/ptodata/2/pubppaa/US08_PUBCOMB pep:*
9: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB pep:*
10: /cgn2_6/ptodata/2/pubppaa/US09_PUBCOMB pep:*
11: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB pep:*
12: /cgn2_6/ptodata/2/pubppaa/US10_PUBCOMB pep:*
13: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB pep:*
14: /cgn2_6/ptodata/2/pubppaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4443	100.0	838	9	US-09-927-315-9
2	4392.5	98.9	839	9	US-09-897-427A-4
3	4392.5	98.9	839	9	US-10-035-045-21
4	3514	79.1	669	9	US-10-096-144-7
5	3514	79.1	669	12	US-10-124-598-7
6	3203.5	72.1	843	9	US-10-096-144-1
7	3203.5	72.1	843	9	US-09-927-315-2
8	3203.5	72.1	843	12	US-10-124-598-1
9	3151.5	70.9	843	9	US-10-096-144-2
10	3151.5	70.9	843	9	US-09-927-315-8
11	3151.5	70.9	843	12	US-10-124-598-2
12	1454	32.7	840	9	US-09-361-652-1
13	1454	32.7	840	9	US-09-927-315-1
14	1434	32.3	842	9	US-09-361-652-2
15	1434	32.3	842	9	US-09-927-315-2
16	1354	30.5	841	9	US-09-897-427A-2
17	1354	30.5	841	9	US-09-799-629-17
18	1354	30.5	841	9	US-10-035-045-17
19	1354	30.5	841	10	US-09-819-946-2

20	1270	28.6	777	9	US-09-361-652-3	Sequence 3, Appl1
21	1270	28.6	777	9	US-09-927-315-3	Sequence 3, Appl1
22	1267.5	28.5	1138	9	US-10-261-482-4	Sequence 4, Appl1
23	1260.5	28.4	763	10	US-09-819-946-4	Sequence 4, Appl1
24	1094	24.6	1078	12	US-10-002-854-2	Sequence 2, Appl1
25	1093	24.6	1078	10	US-09-727-205-2	Sequence 2, Appl1
26	1092.5	24.6	1027	9	US-10-125-792-2	Sequence 2, Appl1
27	1092.5	24.6	1027	9	US-10-125-778-2	Sequence 2, Appl1
28	1078	24.3	1078	9	US-10-125-792-28	Sequence 28, Appl1
29	1078	24.3	1078	9	US-10-125-778-28	Sequence 28, Appl1
30	1056.5	23.8	941	9	US-10-125-792-8	Sequence 8, Appl1
31	1056.5	23.8	941	9	US-10-125-778-8	Sequence 8, Appl1
32	1049.5	23.6	941	9	US-10-125-792-10	Sequence 10, Appl1
33	1049.5	23.6	941	9	US-10-125-778-10	Sequence 10, Appl1
34	1039	23.4	858	9	US-09-799-629-14	Sequence 14, Appl1
35	1039	23.4	858	9	US-09-927-315-25	Sequence 25, Appl1
36	1039	23.4	858	9	US-10-035-045-14	Sequence 14, Appl1
37	1025.5	23.1	850	9	US-10-125-792-12	Sequence 12, Appl1
38	1025.5	23.1	850	9	US-10-125-778-12	Sequence 12, Appl1
39	1014	22.8	852	9	US-10-261-482-2	Sequence 2, Appl1
40	1014	22.8	852	9	US-10-282-837-14	Sequence 14, Appl1
41	1014	22.8	852	10	US-09-796-338A-14	Sequence 14, Appl1
42	1014	22.8	858	9	US-09-927-315-18	Sequence 18, Appl1
43	1013	22.8	858	9	US-09-927-315-20	Sequence 20, Appl1
44	1012	22.8	858	9	US-09-927-315-23	Sequence 23, Appl1
45	1011	22.8	852	9	US-09-897-427A-6	Sequence 6, Appl1

ALIGNMENTS

RESULT 1
US-09-927-315-9
Sequence 9, Application US/09927315
Publication No. US20030040045A1

GENERAL INFORMATION:

APPLICANT: Zuker, Charles S.
APPLICANT: Ryba, Nicholas J.P.

APPLICANT: Nelson, Greg
APPLICANT: Hoon, Mark A.

APPLICANT: Chandrasekar, Jayaram
APPLICANT: Zhang, Yifeng

APPLICANT: The Regents of the University of California
APPLICANT: The Government of the United States of America

APPLICANT: as represented by the Secretary of the
Department of Health and Human Services

TITLE OF INVENTION: Mammalian Sweet Taste Receptors
FILE REFERENCE: 02307E-120110US

CURRENT FILING DATE: 2001-08-10
PRIORITY FILING DATE: 2001-07-03

PRIOR APPLICATION NUMBER: US 60/302,898
NUMBER OF SEQ ID NOS: 25

SOFTWARE: PatentIn ver. 2.1
SEQ ID NO 9

LENGTH: 838
TYPE: PRT

ORGANISM: Homo sapiens
FEATURE:

OTHER INFORMATION: human T1R2 sweet taste receptor
US-09-927-315-9

Query Match 100.0%; Score 4443; DB 9; Length 838;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 838; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGPRKATCSLFFLLWVLAIEPAENSDFYLPDYLGGFSLHANNKGIYHNFLOVPACK 60
DB 1 MGPRKATCSLFFLLWVLAIEPAENSDFYLPDYLGGFSLHANNKGIYHNFLOVPACK 60
QY 61 EYEKVGIVNLMQAMRAFAVEEINNDSSLLPGVLLGAYIVDVCYISNNVQPVLYFLAHEDN 120
DB 61 EYEKVGIVNLMQAMRAFAVEEINNDSSLLPGVLLGAYIVDVCYISNNVQPVLYFLAHEDN 120

QY 121 LPIQEDYSNYSIRVAVIGPDNSESVMYANFLSLFLPQITYSAISDELDRKVRPAL 180
DB 121 LPIQEDYSNYSIRVAVIGPDNSESVMYANFLSLFLPQITYSAISDELDRKVRPAL 180
QY 181 LKTPSADHVEAVOYMLHFRNMWIIYVSSDPTGNDGOLGGERARDDICIAFOETL 240
DB 181 LKTPSADHVEAVOYMLHFRNMWIIYVSSDPTGNDGOLGGERARDDICIAFOETL 240
QY 241 PTLQPNQNMNTSEERORLVTYDKLQOSTARVYVFSPLDLYHFENELRONFTGAWIA 300
DB 241 PTLQPNQNMNTSEERORLVTYDKLQOSTARVYVFSPLDLYHFENELRONFTGAWIA 300
QY 301 SESWAIDPVLANTELGLGFTLIGITIOSVPIPGFSEFERMGPOAGPPPLSRTSOSTCN 360
DB 301 SESWAIDPVLANTELGLGFTLIGITIOSVPIPGFSEFERMGPOAGPPPLSRTSOSTCN 360
QY 361 QECNCLNATLSFNTILRLSEGERVYVSAYAVAAHLSLGCDSCTCKRKYVPMOL 420
DB 361 QECNCLNATLSFNTILRLSEGERVYVSAYAVAAHLSLGCDSCTCKRKYVPMOL 420
QY 421 LEEIMKVNFTLLDHOIFEDPOGVALHLEIYOMQMDRSQNPQSVASYPLQROLKNIQD 480
DB 421 LEEIMKVNFTLLDHOIFEDPOGVALHLEIYOMQMDRSQNPQSVASYPLQROLKNIQD 480
QY 481 ISWHTVNNITIPMSCKRCSGQKRPVGIHCCFECIDCLPGFTLNHTDEVECOACPN 540
DB 481 ISWHTVNNITIPMSCKRCSGQKRPVGIHCCFECIDCLPGFTLNHTDEVECOACPN 540
QY 541 NEMYSOSETSCFRQOLVFLHMEHAPTAVALALAGLSTALIVIMRHOPTIVSAG 600
DB 541 NEMYSOSETSCFRQOLVFLHMEHAPTAVALALAGLSTALIVIMRHOPTIVSAG 600
QY 601 GPMCFMLTLLVAVYVVPYVGPVKVSTCLCROALPCLCTICISCIASVRSQIYCAFK 660
DB 601 GPMCFMLTLLVAVYVVPYVGPVKVSTCLCROALPCLCTICISCIASVRSQIYCAFK 660
QY 661 MASFPRAYSYWRVYQGPYSMAFITVLMKVIVYIGMLARPOSHPRTDDPKITIVSCN 720
DB 661 MASFPRAYSYWRVYQGPYSMAFITVLMKVIVYIGMLARPOSHPRTDDPKITIVSCN 720
QY 721 NPNYRNSLNFNTSLDLLSVGFSAVMGKELPTNYNEAKFTLSMTYFTSSVSLCTFM 780
DB 721 NPNYRNSLNFNTSLDLLSVGFSAVMGKELPTNYNEAKFTLSMTYFTSSVSLCTFM 780
QY 781 AYSGLVTVITVDLTVLNLALISLGYGPKCYMILFPERNTPAYFNSMIOGYTMRD 838
DB 781 AYSGLVTVITVDLTVLNLALISLGYGPKCYMILFPERNTPAYFNSMIOGYTMRD 838

RESULT 2
US-09-897-427A-4
; Sequence 4, Application US/09897427A
; Patent No. US20020160424A1
; GENERAL INFORMATION:
; APPLICANT: ADLER, JON ELLIOT
; APPLICANT: LI, XIADONG
; APPLICANT: STASZEWSKI, LENA
; APPLICANT: XU, HONG
; APPLICANT: EHEVERRT, FERNANDO
; TITLE OF INVENTION: T1R HETERO-OLIGOMERIC TASTE RECEPTORS
; FILE REFERENCE: 078003-0282558
; CURRENT FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 839
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-427A-4

Query Match 98.9%; Score 4392.5; DB 9; Length 839;

Best Local Similarity 99.3%; Pred. No. 0;
Matches 833; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
QY 1 MGPRAKTICSLFLMWIAPAEKSDPYLPQDYLGLGSLHNMKGIYVNLQVPMCK 60
DB 1 MGPRAKTICSLFLMWIAPAEKSDPYLPQDYLGLGSLHNMKGIYVNLQVPMCK 60
QY 61 EYEKVIYGNLMQAMRAVEEINNDSSLPGVLLGYEIVVCYISNNVQVLYFLAHEDN 120
DB 61 EYEKVIYGNLMQAMRAVEEINNDSSLPGVLLGYEIVVCYISNNVQVLYFLAHEDN 120
QY 121 LPIQEDYSNYSIRVAVIGPDNSESVMYANFLSLFLPQITYSAISDELDRKVRPAL 180
DB 121 LPIQEDYSNYSIRVAVIGPDNSESVMYANFLSLFLPQITYSAISDELDRKVRPAL 180
QY 181 LKTPSADHVEAVOYMLHFRNMWIIYVSSDPTGNDGOLGGERARDDICIAFOETL 240
DB 181 LKTPSADHVEAVOYMLHFRNMWIIYVSSDPTGNDGOLGGERARDDICIAFOETL 240
QY 241 PTLQPNQNMNTSEERORLVTYDKLQOSTARVYVFSPLDLYHFENELRONFTGAWIA 300
DB 241 PTLQPNQNMNTSEERORLVTYDKLQOSTARVYVFSPLDLYHFENELRONFTGAWIA 300
QY 301 SESWAIDPVLANTELGLGFTLIGITIOSVPIPGFSEFERMGPOAGPPPLSRTSOSTCN 360
DB 301 SESWAIDPVLANTELGLGFTLIGITIOSVPIPGFSEFERMGPOAGPPPLSRTSOSTCN 360
QY 361 QECNCLNATLSFNTILRLSEGERVYVSAYAVAAHLSLGCDSCTCKRKYVPMOL 420
DB 361 QECNCLNATLSFNTILRLSEGERVYVSAYAVAAHLSLGCDSCTCKRKYVPMOL 420
QY 421 LEEIMKVNFTLLDHOIFEDPOGVALHLEIYOMQMDRSQNPQSVASYPLQROLKNIQD 480
DB 421 LEEIMKVNFTLLDHOIFEDPOGVALHLEIYOMQMDRSQNPQSVASYPLQROLKNIQD 480
QY 481 ISWHTVNNITIPMSCKRCSGQKRPVGIHCCFECIDCLPGFTLNHTDEVECOACPN 540
DB 481 ISWHTVNNITIPMSCKRCSGQKRPVGIHCCFECIDCLPGFTLNHTDEVECOACPN 540
QY 541 NEMYSOSETSCFRQOLVFLHMEHAPTAVALALAGLSTALIVIMRHOPTIVSAG 600
DB 541 NEMYSOSETSCFRQOLVFLHMEHAPTAVALALAGLSTALIVIMRHOPTIVSAG 600
QY 601 GPMCFMLTLLVAVYVVPYVGPVKVSTCLCROALPCLCTICISCIASVRSQIYCAFK 660
DB 601 GPMCFMLTLLVAVYVVPYVGPVKVSTCLCROALPCLCTICISCIASVRSQIYCAFK 660
QY 661 MASFPRAYSYWRVYQGPYSMAFITVLMKVIVYIGMLARPOSHPRTDDPKITIVSC 719
DB 661 MASFPRAYSYWRVYQGPYSMAFITVLMKVIVYIGMLARPOSHPRTDDPKITIVSC 719
QY 720 NPNYRNSLNFNTSLDLLSVGFSAVMGKELPTNYNEAKFTLSMTYFTSSVSLCTFM 779
DB 720 NPNYRNSLNFNTSLDLLSVGFSAVMGKELPTNYNEAKFTLSMTYFTSSVSLCTFM 779
QY 781 AYSGLVTVITVDLTVLNLALISLGYGPKCYMILFPERNTPAYFNSMIOGYTMRD 838
DB 781 AYSGLVTVITVDLTVLNLALISLGYGPKCYMILFPERNTPAYFNSMIOGYTMRD 838

RESULT 3
US-10-035-045-21
; Sequence 21, Application US/10035045
; Publication No. US20030054448A1
; GENERAL INFORMATION:
; APPLICANT: ADLER, JON ELLIOT
; APPLICANT: LI, XIADONG
; APPLICANT: STASZEWSKI, LENA
; APPLICANT: O'CONNELL, SHAWN
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: T1R TASTE RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: 078003-0280681
; CURRENT APPLICATION NUMBER: US/10/035, 045

;; CURRENT FILING DATE: 2002-01-03
;; PRIOR APPLICATION NUMBER: 60/259,227
;; PRIOR FILING DATE: 2001-01-03
;; PRIOR APPLICATION NUMBER: 60/284,547
;; PRIOR FILING DATE: 2001-04-19
;; NUMBER OF SEQ ID NOS: 24
;; SOFTWARE: PatentIn Ver. 2.1
;; SEQ ID NO 21
;; LENGTH: 839
;; TYPE: PRT
;; ORGANISM: Homo sapiens
US-10-035-045-21

Query Match 98.9%; Score 4392.5; DB 9; Length 839;
Best Local Similarity 99.3%; Pred. No. 0;

Matches 833; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

```
QY 1 MGPRAATICSLEFLLWVLAEPANSDPYLPGLDGLGSLHANMKGIYHLNQLQYPMCK 60
DB 1 MGPRAATICSLEFLLWVLAEPANSDPYLPGLDGLGSLHANMKGIYHLNQLQYPMCK 60
QY 61 EYEVKIVGYNLQAMKAFVEEINDDSLPGVLLGYEIVDVCYISNNVQPVLYFLAHEDN 120
DB 61 EYEVKIVGYNLQAMKAFVEEINDDSLPGVLLGYEIVDVCYISNNVQPVLYFLAHEDN 120
QY 121 LPIQEDYNTISRVAIVGIPDNSESMTVANFLSLFLPOLITYSAISDELKRVKVPAL 180
DB 121 LPIQEDYNTISRVAIVGIPDNSESMTVANFLSLFLPOLITYSAISDELKRVKVPAL 180
QY 181 LPTPSADHHEAVQMLHFRMWIIVLVSSDPTVGRDNGOLLGERVARDDICIAFOETL 240
DB 181 LPTPSADHHEAVQMLHFRMWIIVLVSSDPTVGRDNGOLLGERVARDDICIAFOETL 240
QY 241 PTLQPNQNTSEERQRLVTVKLOQSTARVYVESPDLTKHFENEVLRONFTGAWMTA 300
DB 241 PTLQPNQNTSEERQRLVTVKLOQSTARVYVESPDLTKHFENEVLRONFTGAWMTA 300
QY 301 SESMAIDPVLAHLTELHGLGTEFLITIQSVPIPGFSEFERMGPOAPPLSRTSOSTCN 360
DB 301 SESMAIDPVLAHLTELHGLGTEFLITIQSVPIPGFSEFERMGPOAPPLSRTSOSTCN 360
QY 361 QECDCNLANLTSFNTILRLSGERVYVSVAAYAAHALHSLGCDKSTCTKRVYPMOL 420
DB 361 QECDCNLANLTSFNTILRLSGERVYVSVAAYAAHALHSLGCDKSTCTKRVYPMOL 420
QY 421 LBEIMKVNFTLDHOFEPDQGVLAHLFIYOMQMDRSONPQSVASVYPLQRLKNID 480
DB 421 LBEIMKVNFTLDHOFEPDQGVLAHLFIYOMQMDRSONPQSVASVYPLQRLKNID 480
QY 481 ISMHTVNNITIPMSCKRQSGOKKRPVGIHVCCFECIDCLPGLTFLNHTEDEECQACPN 540
DB 481 ISMHTVNNITIPMSCKRQSGOKKRPVGIHVCCFECIDCLPGLTFLNHTEDEECQACPN 540
QY 541 NEMSVQSESCFRQLVFLMEHEAPITAVALLAAGLSTLAILVTFMRHFOPTVRSAG 600
DB 541 NEMSVQSESCFRQLVFLMEHEAPITAVALLAAGLSTLAILVTFMRHFOPTVRSAG 600
QY 601 GPMCFMLTLLVAVYVYVGPVSTCLROALFPLCFITICISCIIVRSQIYCAK 660
DB 601 GPMCFMLTLLVAVYVYVGPVSTCLROALFPLCFITICISCIIVRSQIYCAK 660
QY 661 MASRPRAVSVMYRGYPVSMATITVLMKVIIVIGMLARPOS-HPRTPDDPKITIVSC 719
DB 661 MASRPRAVSVMYRGYPVSMATITVLMKVIIVIGMLARPOS-HPRTPDDPKITIVSC 719
QY 720 NENYRNSLTFNLSLDSVGSFAYMGKELPTYNNEAKFTTSLNTEFTTSVSLSCTEM 779
DB 720 NENYRNSLTFNLSLDSVGSFAYMGKELPTYNNEAKFTTSLNTEFTTSVSLSCTEM 779
QY 780 SAYSGVLTIVDLTVLNLALISLGYFGPKCYMLTFYERHPTPAVFNMIQGYMRD 838
DB 780 SAYSGVLTIVDLTVLNLALISLGYFGPKCYMLTFYERHPTPAVFNMIQGYMRD 838
```

RESULT 4
US-10-096-144-7
Sequence 7, Application US/10096144
Publication No. US20030022288A1
GENERAL INFORMATION:

;; APPLICANT: Zuker, Charles S.
;; APPLICANT: Adler, Jon Elliott
;; APPLICANT: Lindemeier, Juergen
;; APPLICANT: The Regents of the University of California
;; TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
;; TITLE OF INVENTION: Involved in Sensory Transduction
;; FILE REFERENCE: 02307E-088720US
;; CURRENT APPLICATION NUMBER: US/10/096,144

;; PRIOR APPLICATION NUMBER: 09/361,631
;; PRIOR FILING DATE: 2002-03-08
;; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-27
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/095,464
;; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-28
;; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/112,747

;; NUMBER OF SEQ ID NOS: 10
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 7
;; LENGTH: 669
;; TYPE: PRT
;; ORGANISM: Homo sapiens

;; OTHER INFORMATION: human G-protein coupled receptor (GPCR) B4 amino
;; OTHER INFORMATION: acid sequence
US-10-096-144-7

Query Match 79.1%; Score 3514; DB 9; Length 669;
Best Local Similarity 98.4%; Pred. No. 3,6e-288;

Matches 666; Conservative 1; Mismatches 2; Indels 8; Gaps 2;

```
QY 162 ITYSASIDELRDKVRFPALRTTPSADHHEAVQMLHFRMWIIVLVSSDPTVGRDNG 221
DB 1 ITYSASIDELRDKVRFPALRTTPSADHHEAVQMLHFRMWIIVLVSSDPTVGRDNG 221
QY 222 LGGERVARDDICIAFOETLPTLQPNQNTSEERQRLVTVKLOQSTARVYVESPDLTK 281
DB 222 LGGERVARDDICIAFOETLPTLQPNQNTSEERQRLVTVKLOQSTARVYVESPDLTK 281
QY 282 YHFEENEVLRONFTGAWMTASESWAIDPVLAHLTELHGLGTEFLITIQSVPIPGFSE 341
DB 282 YHFEENEVLRONFTGAWMTASESWAIDPVLAHLTELHGLGTEFLITIQSVPIPGFSE 341
QY 342 GPOAGPPPLSRTSOSTYCNQECDCNLANLTSFNTILRLSGERVYVSVAAYAAHALS 401
DB 342 GPOAGPPPLSRTSOSTYCNQECDCNLANLTSFNTILRLSGERVYVSVAAYAAHALS 401
QY 402 LIGCDKSTCTKRVYPMOLLEIIMKVNFTLDHOFEPDQGVLAHLFIYOMQMDRSONP 461
DB 402 LIGCDKSTCTKRVYPMOLLEIIMKVNFTLDHOFEPDQGVLAHLFIYOMQMDRSONP 461
QY 461 POSVASYPLQRLKNIK-TSLHTVNNITIPMSCKRQSGOKKRPVGIHVCCFECIDCL 521
DB 461 POSVASYPLQRLKNIK-TSLHTVNNITIPMSCKRQSGOKKRPVGIHVCCFECIDCL 521
QY 522 PGTFNLNHTDEVEEQACPNNEMSVQSESCFRQLVFLMEHEAPITAVALLAAGLSTL 581
DB 522 PGTFNLNHTDEVEEQACPNNEMSVQSESCFRQLVFLMEHEAPITAVALLAAGLSTL 581
QY 582 ALIVTFMRHFOPTVRSAGGPMCFMLTLLVAVYVYVGPVSTCLROALFPLCF 641
DB 582 ALIVTFMRHFOPTVRSAGGPMCFMLTLLVAVYVYVGPVSTCLROALFPLCF 641
QY 642 TTCISCIIVRSQIYCAFMASRPRAVSVMYRGYPVSMATITVLMKVIIVIGMLAR 701
DB 642 TTCISCIIVRSQIYCAFMASRPRAVSVMYRGYPVSMATITVLMKVIIVIGMLAR 701
QY 702 QSHPRTPDDPKITIVSCPNYRNSLTFNLSLDSVGSFAYMGKELPTYNNEAKFT 761
```

```

|||||
Db 533 QSHRPDPDPKTIIVSCNPNYRNSLFFNTSLDLLSVGFSPAYMKELPTYNNEAKFI 592
QY 762 TLSMTEFFTSVSLCTFMSAYSGVLTIVDLTVLNLALISLGYEGPKCYMILFPERN 821
Db 593 TLSMTEFFTSVSLCTFMSAYSGVLTIVDLTVLNLALISLGYEGPKCYMILFPERN 652
QY 822 TPAYFNSMIGYTMRRD 838
Db 653 TPAYFNSMIGYTMRRD 669

```

```

RESULT 5
US-10-124-598-7
; Sequence 7, Application US/10124598
; Patent No. US20020119526A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Lindemeyer, Juerger
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
; TITLE OF INVENTION: Involved in Sensory Transduction
; FILE REFERENCE: 02307E-088720US
; CURRENT APPLICATION NUMBER: US/10/124, 598
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/361,631
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/112,747
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human G-protein coupled receptor (GPCR) B4 amino
; OTHER INFORMATION: acid sequence
US-10-124-598-7

```

Query Match 79.1%; Score 3514; DB 12; Length 669;
 Best Local Similarity 98.4%; Pred. No. 3.6e-288;
 Matches 666; Conservative 1; Mismatches 2; Indels 8; Gaps 2;

```

QY 162 ITYSALSDERLDRKVRFPALRTTPSADHVEAMVOLMHRMNMIIYLVSSDYGRNGQ 221
Db 1 ITYSALSDERLDRKVRFPALRTTPSADHVEAMVOLMHRMNMIIYLVSSDYGRNGQ 60
QY 222 LIGERVARRDICIAFOETLPTLOPNOMTSEEROLVTIVDKLQOSTARVYVFPDLTL 281
Db 61 LIGERVARRDICIAFOETLPTLOPNOMTSEEROLVTIVDKLQOSTARVYVFPDLTL 120
QY 282 YHFEENEVLDRONFTGAWIASMAIDVYLNHLELGHGTFGLTIGTISVPIPGSEEREW 341
Db 121 YHFEENEVLDRONFTGAWIASMAIDVYLNHLELGHGTFGLTIGTISVPIPGSEEREW 180
QY 342 GPQAGPPPLSRSTOSYTCNOCNCLNATLSFNTILRLSERVYVSYSAVYAAHLS 401
Db 181 GPQAGPPPLSRSTOSYTCNOCNCLNATLSFNTILRLSERVYVSYSAVYAAHLS 240
QY 402 LIGCDKSTCKRVVYPMQLLEIWKVNFLLDHOIFFDPQGDVALHLEIYQWMDRSQNP 461
Db 241 LIGCDKSTCKRVVYPMQLLEIWKVNFLLDHOIFFDPQGDVALHLEIYQWMDRSQNP 300
QY 462 FQSVASYTPYQROLKNTQDISMTVNTTMSKCSKRCQSGQKKRPVGIHCCBECIDCL 521
Db 301 FQSVASYTPYQROLKNTK -TSLHTVNTTIPMSKCSKRCQSGQKKRPVGIHCCBECIDCL 359
QY 522 PGTEFLNTEDEYEQACPNMWSYQSTSCFKRLVFLWHEAPTIVALLAALGFI STL 581
Db 360 PGTEFLNTEDEYEQACPNMWSYQSTSCFKRLVFLWHEAPTIVALLAALGFI STL 412

```

```

QY 582 AILVIFWRHFDPIVNSAGPMCFMLTLLVAIVMVVYVGPVKSTCLCRQALPELPCF 641
Db 413 AILVIFWRHFDPIVNSAGPMCFMLTLLVAIVMVVYVGPVKSTCLCRQALPELPCF 472
QY 642 TICISCIANVSFOIVCAFKNASRPFRAYSYWYVQGYVSMATITVYKMTIYIGMLARP 701
Db 473 TICISCIANVSFOIVCAFKNASRPFRAYSYWYVQGYVSMATITVYKMTIYIGMLARP 532
QY 702 QSHRPDPDPKTIIVSCNPNYRNSLFFNTSLDLLSVGFSPAYMKELPTYNNEAKFI 761
Db 533 QSHRPDPDPKTIIVSCNPNYRNSLFFNTSLDLLSVGFSPAYMKELPTYNNEAKFI 592
QY 762 TLSMTEFFTSVSLCTFMSAYSGVLTIVDLTVLNLALISLGYEGPKCYMILFPERN 821
Db 593 TLSMTEFFTSVSLCTFMSAYSGVLTIVDLTVLNLALISLGYEGPKCYMILFPERN 652
QY 822 TPAYFNSMIGYTMRRD 838
Db 653 TPAYFNSMIGYTMRRD 669

```

```

RESULT 6
US-10-096-144-1
; Sequence 1, Application US/10096144
; Publication No. US2003002288A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Lindemeyer, Juerger
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
; TITLE OF INVENTION: Involved in Sensory Transduction
; FILE REFERENCE: 02307E-088720US
; CURRENT APPLICATION NUMBER: US/10/096,144
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/361,631
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/095,464
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/112,747
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.0
; LENGTH: 843
; TYPE: PRT
; ORGANISM: Rattus sp.
; FEATURE:
; OTHER INFORMATION: rat G-protein coupled receptor (GPCR) B4 amino
; OTHER INFORMATION: acid sequence
US-10-096-144-1

```

Query Match 72.1%; Score 3203.5; DB 9; Length 843;
 Best Local Similarity 70.6%; Pred. No. 8.3e-262;
 Matches 595; Conservative 108; Mismatches 133; Indels 7; Gaps 4;

```

QY 1 MGPRAKTICSLFLLWYLAEP--AENSDFYLPQDYLLGSLSHANMKGIVHLNFIQVP 57
Db 1 MGPRAKTICSLFLLWYLAEP--AENSDFYLPQDYLLGSLSHANMKGIVHLNFIQVP 60
QY 58 MCKEYEVKIVGINLMQAMRAVEEINNDSSILBGLVLYGELYDVYCYISNNVQPLTYFLAH 117
Db 61 MCKEYEVKIVGINLMQAMRAVEEINNDSSILBGLVLYGELYDVYCYISNNVQPLTYFLAH 120
QY 118 EDNLPLTQEDYSNTISVNVAVIGPDNSESVMYVNFSLFLPQITYSALSDERLDRKVR 177
Db 121 DDLLPLTKDYSOYMPHVAVVIGPDNSESATVSNLSHFLPQITYSALSDERLDRKVR 180
QY 178 PALIRTPSADHVEAMVOLMHRMNMIIYLVSSDYGRNGQDLGERVAF-RDICIACF 236
Db 181 PSMKRTVPSATHTHEANVOLMHRMNMIIYLVSSDYGRNGQDLGERVAF-RDICIACF 240
QY 237 QETLPTLOPNOMTSEEROLVTIVDKLQOSTARVYVFPDLTLVYHFEENEVLDRONFTGA 296

```

Db	241	OEVLPIRESSQVMRSEEQROLDNIDJLKRTRSAWVYVFEPLSIVSFFHEVLWRMNTGF	3000
Qy	297	VWIASSEWALDPVYHNHTELGHTFLGCTITQSYVPJGSESEFRMGCPQACBPPLSRTSOS	3366
Db	301	VWIASSEWALDPVYHNHTELGHTGTELGVIQIORSIPGFSOFRRRKRDPYVPVNTNLR	3600
Qy	357	YTCNOECDNCINATLSEFWILRLSGEVRVYSVSAVVAVAHAAHLSLJGCDKSYCTKRHY	4166
Db	361	YTCNOCDACJLNTKSFNNILISGEVRVYSVSAVVAVAHAAHLLRLLGCRVACTKQYV	4200
Qy	417	PWOLLEELWAKVYFPLLDHQIFEDPQGVALHLEIVOMQDRSONFQOSVASYFPLQRLK	4766
Db	421	PWOLIRELWYVNFLLGNRLEFFDQGGMPMLDIIOWMDLSONFQOSIASYFSTRKLT	4800
Qy	477	NIODISMTWVNTLPMSCSCSRQSGGKKRPVGIHCCFPCIDLPSTGFANHEDDECO	5366
Db	481	YINNVSWYTPNNYTPVEMCSISQCPGQMKMSVGLHPCCFECDCMBPTYLNRSDENCL	5400
Qy	537	ACPNNEWSYQSETCFRQVLFEWHEAPTTAVALLAAGLSTLALIVYFWRHFOPIY	5966
Db	541	SCPSPMSYKINDICFOFRPFLEHMEHVPTIYVAILAALGFESTYALIFTFWRHFEQPMV	6000
Qy	597	RSAGCBPCFLMTLLLYAAVYVNPVYVBPFRVSYTCLCQALFPLCTYICICIAVRSQIV	6566
Db	601	RSAGGBCPCLMTLPLLLAFGAVPYVBPPTVFCSCQDAEFVYCSCILCITIVRSQIV	6600
Qy	657	CAEFMASRFPRAVSUYWRYQGPVYSMAFIVLWLVIVVIMLARPOSH--RTDDPPKI	7144
Db	661	CVFEMARRPLSAVSFWMRHGPYFVAFIRAIKALVVGMLA--TTINPIGRIDDDPNI	7190
Qy	715	TVYSCNNYRNSLLEFNTSLDLLISVGSFSAIYMGKEPLTYNNEAKFTTLSMTYFPSSVS	7744
Db	720	MILSCHNRYRNGLEFNTSMOLLISVLGFSAPYMGKEPLTYNNEAKFTTLSMTSESTSSIS	7790
Qy	775	LCTFMSASGVLYTVLDLVTVLNLALISIGYGPCKMYLLEPERNTPYFNSMIGYV	8344
Db	780	LCTFMSVHGDVLYTMDLVTVLNFALIGYGPCKMYLLEPERNTPYFNSMIGYV	8390
Qy	835	MRR 837.	
Db	840	MRR 842	

RESULT 7
 US-09-927-315-7
 Sequence 7, Application US/09927315
 Publication No. US20030040045A1
 GENERAL INFORMATION:
 APPLICANT: Zuker, Charles S.
 APPLICANT: Ryba, Nicholas J. P.
 APPLICANT: Nelson, Greg
 APPLICANT: Hoon, Mark A.
 APPLICANT: Chandrasekar, Jayaram
 APPLICANT: Zhang, Yifeng
 APPLICANT: The Regents of the University of California
 APPLICANT: The Government of the United States of America
 APPLICANT: as represented by the Secretary of the
 APPLICANT: Department of Health and Human Services
 TITLE OF INVENTION: Mammalian Sweet Taste Receptors
 FILE REFERENCE: 02307E-120110US
 CURRENT APPLICATION NUMBER: US/09/927, 315
 CURRENT FILING DATE: 2001-08-10
 PRIOR APPLICATION NUMBER: US 60/302,898
 PRIOR FILING DATE: 2001-07-03
 NUMBER OF SEQ ID NOS: 25
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 7
 LENGTH: 843
 TYPE: PRT
 ORGANISM: Rattus sp.
 FEATURE:
 OTHER INFORMATION: rat Tir2 sweet taste receptor

US-09-927-315-7

Query Match	72.1%;	Score 3203.5;	DB 9;	Length 843;
Best Local Similarity	70.6%;	Pred. No. 8.3e-252;		
Matches 595;	Conservative 108;	Mismatches 133;	Indels 7;	Gaps 4

[illegible]

```

RESULT 8
US-10-124-598-1
; Sequence 1, Application US/10124598
; Patent No. US20020119526A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.

```

```

; APPLICANT: Adler, Jon Elliot
; APPLICANT: Lindemeyer, Juergen
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
; TITLE OF INVENTION: Involved in Sensory Transduction
; FILE REFERENCE: 02307E-08872005
; CURRENT APPLICATION NUMBER: US/10/124,598
; PRIOR FILING DATE: 2002-04-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/361,631
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/112,747
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 1
; LENGTH: 843
; TYPE: PRN
; ORGANISM: Rattus sp.
; FEATURE:
; OTHER INFORMATION: rat G-protein coupled receptor (GPCR) B4 amino
; OTHER INFORMATION: acid sequence
; US-10-124-598-1

```

```

Query Match          72.1%; Score 3203.5; DB 12; Length 843;
Best Local Similarity 70.6%; Pred. No. 8.3e-262;
Matches 595; Conservative 108; Mismatches 133; Indels 7; Gaps 4;

```

```

QY 1 MGPARTICSLFFLLWLAEP---AENSDFYLPDGYLLGGLFSLHANKGIYHLNFOVP 57
DB 1 MGPARTICSLFLLWLAEP---AENSDFYLPDGYLLGGLFSLHANKGIYHLNFOVP 60
QY 58 MCKEYKAVIGYINLMQAMRAVEEINNDSSLLPGVLLGYEIVDYCYISNNVQVLYFLAH 117
DB 61 KCNEFTKAVIGYINLMQAMRAVEEINNDSSLLPGVLLGYEIVDYCYISNNVQVLYFLAH 120
QY 118 EDNLPLTQEOYSNYSISVAVIIGPDSSEVMTANFLSLFLPQITYSAISDELDRDYRF 177
DB 121 DDLLPLTKDYSQYMPHVAIVIGPDSSEVMTANFLSLFLPQITYSAISDELDRDYRF 180
QY 178 PALRTTPSADHVEAVOVLALHFRWNIIVLVSSDYGNDQGLGERVAR-RDICIATF 236
DB 181 PSMRTVPSAHTHEAVOVLALHFRWNIIVLVSSDYGNDQGLGERVAR-RDICIATF 240
QY 237 QETPLTQPNOMNTSEERQRLVTVDKLQOSTARVAVVFPDLYLHFFNEVLRQNTGA 296
DB 241 QETPLTQPNOMNTSEERQRLVTVDKLQOSTARVAVVFPDLYLHFFNEVLRQNTGA 300
QY 297 VWIASEMAIDPVLAHNLTELHGTFLGVTIQRSVIFGFSFREMGPQAGPPPLSRISQS 356
DB 301 VWIASEMAIDPVLAHNLTELHGTFLGVTIQRSVIFGFSFREMGPQAGPPPLSRISQS 360
QY 357 YTCNOECDNCINATLSEFNTILRLSGERVVSVYSAYAVAAHLSLGCDSCTCRKRVY 416
DB 361 YTCNOECDNCINATLSEFNTILRLSGERVVSVYSAYAVAAHLSLGCDSCTCRKRVY 420
QY 417 PMOLLEIKWVNTFLDHOJFFDPQGVALLHLEIVQMDRSONDFQSVASYPLQROLK 476
DB 421 PMOLLEIKWVNTFLDHOJFFDPQGVALLHLEIVQMDRSONDFQSVASYPLQROLK 480
QY 477 NIODISHWNTVTPMSCKRSCOSGOKRRKPVGIHVCCECTIDCLPGFLNHTDEDECQ 536
DB 481 YINNVSWTTPNNTVPMSCSKSCOPGOMKRSVGLHPCCEFLDQMPGTILNRSADENCL 540
QY 537 ACNNENSYOSETSCFRQVLEFMEHAPTLAVALAAGLSTLALIVTFWRHFOPTIV 596
DB 541 SCPSMSYKNDITCFQORPTEFMEHAPTLAVALAAGLSTLALIVTFWRHFOPTIV 600
QY 597 RSAGGPCFMTLLIYLAIVAVVYVGPVKYSTLCRQALPFLCFTICISIAVRSQIV 656
DB 601 RSAGGPCFMTLLIYLAIVAVVYVGPVKYSTLCRQALPFLCFTICISIAVRSQIV 660
QY 657 CAFEMASRPRAYSYWRVYQGPVSMATVYLVKVIIVVIGLARPOSHP--RTDPDPKI 714
DB 657 CAFEMASRPRAYSYWRVYQGPVSMATVYLVKVIIVVIGLARPOSHP--RTDPDPKI 714

```

```

DB 661 CVFMAIRLPSAYSEFMNRHGPVYVAFITAKVALVGNMIA--TTINIGRIDDDPNI 719
QY 715 TIVSCNPNYRSILFNTSLDLSVGFSPAYMKEPLTNYNEAKFTTILSMTEFFTSVS 774
DB 720 MTLSCNPNYRSILFNTSLDLSVGFSPAYMKEPLTNYNEAKFTTILSMTEFFTSVS 779
QY 775 LCTFMSAYSGVLTIVDLVTLNLALISLGFPGKCYMILFPERNTPAVFNSMIGYT 834
DB 780 LCTFMSAYSGVLTIVDLVTLNLALISLGFPGKCYMILFPERNTPAVFNSMIGYT 839
QY 835 MRK 837
DB 840 MRK 842

```

```

RESULT 9
US-10-096-144-2
; Sequence 2, Application US/10096144
; Publication No. US2003002288A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Lindemeyer, Juergen
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
; TITLE OF INVENTION: Involved in Sensory Transduction
; FILE REFERENCE: 02307E-08872005
; CURRENT APPLICATION NUMBER: US/10/096,144
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/361,631
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/095,464
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/112,747
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 2
; LENGTH: 843
; TYPE: PRN
; ORGANISM: Mus sp.
; FEATURE:
; OTHER INFORMATION: mouse G-protein coupled receptor (GPCR) B4 amino
; OTHER INFORMATION: acid sequence
; US-10-096-144-2

```

```

Query Match          70.9%; Score 3151.5; DB 9; Length 843;
Best Local Similarity 68.9%; Pred. No. 2e-257;
Matches 581; Conservative 113; Mismatches 142; Indels 7; Gaps 4;

```

```

QY 1 MGPARTICSLFFLLWLAEP---AENSDFYLPDGYLLGGLFSLHANKGIYHLNFOVP 57
DB 1 MGPARTICSLFLLWLAEP---AENSDFYLPDGYLLGGLFSLHANKGIYHLNFOVP 60
QY 58 MCKEYKAVIGYINLMQAMRAVEEINNDSSLLPGVLLGYEIVDYCYISNNVQVLYFLAH 117
DB 61 KCNEFTKAVIGYINLMQAMRAVEEINNDSSLLPGVLLGYEIVDYCYISNNVQVLYFLAH 120
QY 118 EDNLPLTQEOYSNYSISVAVIIGPDSSEVMTANFLSLFLPQITYSAISDELDRDYRF 177
DB 121 DDLLPLTKDYSQYMPHVAIVIGPDSSEVMTANFLSLFLPQITYSAISDELDRDYRF 180
QY 178 PALRTTPSADHVEAVOVLALHFRWNIIVLVSSDYGNDQGLGERVAR-RDICIATF 236
DB 181 PSMRTVPSAHTHEAVOVLALHFRWNIIVLVSSDYGNDQGLGERVAR-RDICIATF 240
QY 237 QETPLTQPNOMNTSEERQRLVTVDKLQOSTARVAVVFPDLYLHFFNEVLRQNTGA 296
DB 241 QETPLTQPNOMNTSEERQRLVTVDKLQOSTARVAVVFPDLYLHFFNEVLRQNTGA 300
QY 297 VWIASEMAIDPVLAHNLTELHGTFLGVTIQRSVIFGFSFREMGPQAGPPPLSRISQS 356
DB 301 VWIASEMAIDPVLAHNLTELHGTFLGVTIQRSVIFGFSFREMGPQAGPPPLSRISQS 360

```


Best Local Similarity 39.6%; Pred. No. 2.3e-112;
Matches 327; Conservative 123; Mismatches 340; Indels 36; Gaps 15;

```

27 FYLPGDYLGLGSLFSLHNMKGIVLNFLOVPMC-KEYEVKIGYNLMQAMFEAVEEIND 85
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 33 FSLPGDYLGLGSLFSLHNMKGIVLNFLOVPMC-KEYEVKIGYNLMQAMFEAVEEIND 90
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 86 SSLPGVLLGEIYDVCYISNNQPVLYFLAHE-DNLLP10EDYSNTISRVAVIGPDS 144
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 91 TALLPNTTIGELYDVCSESSNYATLRPAQGTGHEMQRLRNSSKVALIGDNT 150
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 145 ESWTVANFLSLFLPQITYSATISDELDRKVRFPALLRTTPSADHNYEAMVQLMHRWN 204
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 151 DHAVTTAALSLPFLMPLVSEASSVILSGKRKFPFLRTIPSDRYQVEVIRLLQSGFW 210
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 205 WIIVLVSDTYGRNGQLGERVARRDICIAFOETPLTLPQNMTESEERQRLTYIDKL 264
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 211 WISLVSGYDGOGLGVQALELAPRGICVAFKDVPL---SQAQDPRMQRML---RL 264
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 265 QOSTARVVVVFSPDLTYHFEENEYLRONFTGAVWIASESNAIDVNLNTELGHLGFTLG 324
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 265 ARARTTVVVVFSNRHLGAVFERSVVLNLTKGVWIASEDMASITTYINVPQIGIGTVLG 324
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 325 ITTOSVPIPGSEFERENGPOA--GPPLSRTSOSTYCNQEDCNLNTLSNTLRISGE 382
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 325 VALQROVPGLKEFEESYVQAVMGAPRTCPREGSNGCTNOJCRechAFTTNMPELAFSM 384
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 383 RVVSVYSAYAVAHALSHLSDGCKSTCKRVVYPMOLLEIMKVNFTLDHOFEPDQ 442
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 385 SAATVNEAYAVAHGHLGCTSGTCARGPVYPMOLLQOITYVNLHKKTYAFDDKG 444
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 443 DVALHLEIVOMQDRSONPQSV--ASYYPLOQRLKNIODISWHTVNTTIPMSCKSRQ 500
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 445 DPLGYDYLIDAMDMNGPMTPEVEIGSASLSPVHLDI-NKTKIQMHGKNNQVAVSCTRCL 503
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 501 SGOKKRVVGHVCCFECIDCLPGFTLNHTEDEYECQACPNMNSYQSETSCFRQVLFLE 560
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 504 EGHRLWVGHSHCCFECMPCEAGTFELN-TSELHQCQCGTEEMAPRESSACFSRTVEFLG 562
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 561 WHEAPTAVALALAGSLTAIVIF-----WRHFQPIYVSAGGPMCFMLTLLV 613
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 563 WHE--PISIVLTA-----NTLLILLIGTAGLFAFR-LHPPVVSAGGRCLFMLGSLVA 615
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 614 AYVWVYVGPVKVSTCLCQALFPLCTTICISIAVRSFOICAFKMASFRPAYSIVW 673
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 616 GSCSLYFPGKPYVACILKQPLFSLGFAIFLSCTLRISQVLIIFRFSKVPFTFYHTWA 675
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 674 RYOGPIYVMAFIVLKNVIVIGMLARQSHPRDTP--DPKITYVSCNPNYNSILFN 730
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 676 QNHGAGI---FVIYVSTVHFLCLTWMAMTPTREYQRFPHLVILCETEVNSVGFIVA 732
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 731 TSIDLILSVGFSFAYMGKELPTNYNEAKFTLISMTFFYSSVSLCTFMSAVGLVTVI 790
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 733 FAHNTLISIFVCSYIGKELPENYNEAKCVTSLLHAFWSIAFTMSSIQSGSYLPAV 792
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 791 DLATVYNLAISIGFPGKCYMLTFPERNTPAFVNSMIOGTMR 836
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 793 NVLAGLTLISGGSFYFLPKCYVILCRPELNTNTEHFOASIODYTRR 838
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

RESULT 15
US-09-927-315-2

; Sequence 2, Application US/09927315
; Publication No. US20030040045A1

; GENERAL INFORMATION:

; APPLICANT: Zuker, Charles S.
; APPLICANT: Ryba, Nicholas J.P.

; APPLICANT: Nelson, Greg
; APPLICANT: Hoon, Mark A.

; APPLICANT: Chandrasekar, Jayaram
; APPLICANT: Zhang, Yifeng

; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America

```

; APPLICANT: as represented by the Secretary of the
; APPLICANT: Department of Health and Human Services
; TITLE OF INVENTION: Mammalian Sweet Taste Receptors
; FILE REFERENCE: 02307E-120110US
; CURRENT APPLICATION NUMBER: US/09/927, 315
; PRIOR FILING DATE: 2001-08-10
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 842
; TYPE: PRT
; ORGANISM: Mus musculus
; FEATURE:
; OTHER INFORMATION: mouse Tr1r1 sweet taste receptor
US-09-927-315-2

```

Query Match

Best Local Similarity 32.3%; Score 1434; DB 9; Length 842;

Matches 327; Conservative 123; Mismatches 340; Indels 36; Gaps 15;

```

27 FYLPGDYLGLGSLFSLHNMKGIVLNFLOVPMC-KEYEVKIGYNLMQAMFEAVEEIND 85
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 33 FSLPGDYLGLGSLFSLHNMKGIVLNFLOVPMC-KEYEVKIGYNLMQAMFEAVEEIND 90
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 86 SSLPGVLLGEIYDVCYISNNQPVLYFLAHE-DNLLP10EDYSNTISRVAVIGPDS 144
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 91 TALLPNTTIGELYDVCSESSNYATLRPAQGTGHEMQRLRNSSKVALIGDNT 150
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 145 ESWTVANFLSLFLPQITYSATISDELDRKVRFPALLRTTPSADHNYEAMVQLMHRWN 204
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 151 DHAVTTAALSLPFLMPLVSEASSVILSGKRKFPFLRTIPSDRYQVEVIRLLQSGFW 210
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 205 WIIVLVSDTYGRNGQLGERVARRDICIAFOETPLTLPQNMTESEERQRLTYIDKL 264
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 211 WISLVSGYDGOGLGVQALELAPRGICVAFKDVPL---SQAQDPRMQRML---RL 264
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 265 QOSTARVVVVFSPDLTYHFEENEYLRONFTGAVWIASESNAIDVNLNTELGHLGFTLG 324
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 265 ARARTTVVVFSNRHLGAVFERSVVLNLTKGVWIASEDMASITTYINVPQIGIGTVLG 324
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 325 ITTOSVPIPGSEFERENGPOA--GPPLSRTSOSTYCNQEDCNLNTLSNTLRISGE 382
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 325 VALQROVPGLKEFEESYVQAVMGAPRTCPREGSNGCTNOJCRechAFTTNMPELAFSM 384
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 383 RVVSVYSAYAVAHALSHLSDGCKSTCKRVVYPMOLLEIMKVNFTLDHOFEPDQ 442
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 385 SAATVNEAYAVAHGHLGCTSGTCARGPVYPMOLLQOITYVNLHKKTYAFDDKG 444
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 443 DVALHLEIVOMQDRSONPQSV--ASYYPLOQRLKNIODISWHTVNTTIPMSCKSRQ 500
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 445 DPLGYDYLIDAMDMNGPMTPEVEIGSASLSPVHLDI-NKTKIQMHGKNNQVAVSCTRCL 503
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 501 SGOKKRVVGHVCCFECIDCLPGFTLNHTEDEYECQACPNMNSYQSETSCFRQVLFLE 560
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 504 EGHRLWVGHSHCCFECMPCEAGTFELN-TSELHQCQCGTEEMAPRESSACFSRTVEFLG 562
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 561 WHEAPTAVALALAGSLTAIVIF-----WRHFQPIYVSAGGPMCFMLTLLV 613
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 563 WHE--PISIVLTA-----NTLLILLIGTAGLFAFR-LHPPVVSAGGRCLFMLGSLVA 615
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 614 AYVWVYVGPVKVSTCLCQALFPLCTTICISIAVRSFOICAFKMASFRPAYSIVW 673
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 616 GSCSLYFPGKPYVACILKQPLFSLGFAIFLSCTLRISQVLIIFRFSKVPFTFYHTWA 675
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 674 RYOGPIYVMAFIVLKNVIVIGMLARQSHPRDTP--DPKITYVSCNPNYNSILFN 730
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 676 QNHGAGI---FVIYVSTVHFLCLTWMAMTPTREYQRFPHLVILCETEVNSVGFIVA 732
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
QY 731 TSIDLILSVGFSFAYMGKELPTNYNEAKFTLISMTFFYSSVSLCTFMSAVGLVTVI 790
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
Db 733 FAHNTLISIFVCSYIGKELPENYNEAKCVTSLLHAFWSIAFTMSSIQSGSYLPAV 792
   :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||

```

OY 791 DLVTVNLALISLGYFGPKCYMLFYPERNTPAYENSMIOGYMR 836
Db 793 NVLAGIATLSGFGSGYFLPKCYIILCRPELNNTTEHFQASIODYTRR 838

Search completed: May 19, 2003, 10:04:41
job time : 43.1769 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 23, 2003, 11:21:44 ; Search time 5176.99 Seconds

(Without alignments)
11299.347 Million cell updates/sec

Title: US-09-927-315-12

Perfect score: 2010

Sequence: 1 atcacctacagcgcacatcag.....accacatgagggagactag 2010

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenBank:*

1: gb_da:*

2: gb_htg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_pln:*

35: em_htg_rnd:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2010	100.0	2010	6 AR208617	AR208617 Sequence
2	1232.8	61.3	2993	6 AR208615	AR208615 Sequence
3	1230	61.2	2529	10 AF127390	AF127390 Rattus no
4	1220.6	60.7	3060	10 AF337041	AF337041 Mus muscu
5	1212.6	60.3	2532	6 AR208616	AR208616 Sequence
6	1212.6	60.3	2532	10 AY032623	AY032623 Mus muscu
7	879.8	43.8	930	2 AF458149S6	AF458149S6 Homo sapi
8	878.2	43.7	209676	2 AL831755	AL831755 Homo sapi
9	774	38.5	774	9 AF458149S3	AF458149S3 Homo sapi
10	605	30.1	206074	2 AL831790	AL831790 Mus muscu
11	474.6	23.6	154891	2 AC115397	AC115397 Rattus no
12	392.8	19.5	2520	10 AF127389	AF127389 Rattus no
13	386.4	19.2	2292	6 AX282894	AX282894 Sequence
14	386.4	19.2	2526	6 AX282892	AX282892 Sequence
15	386.4	19.2	2951	6 AX282900	AX282900 Sequence
16	385.2	19.2	2759	6 AX354027	AX354027 Sequence
17	384.8	19.1	2516	9 AF387618	AF387618 Homo sapi
18	367.2	18.3	2526	10 AF301162	AF301162 Mus muscu
19	359.2	17.9	2529	10 AY032622	AY032622 Mus muscu
20	359.2	17.9	2892	10 AF337040	AF337040 Mus muscu
21	357.6	17.8	2526	10 AF301161	AF301161 Mus muscu
22	258	12.8	697	6 AX320494	AX320494 Sequence
23	218.6	10.9	3458	6 AX417279	AX417279 Sequence
24	217	10.8	2559	6 AX239649	AX239649 Sequence
25	217	10.8	2559	6 AX298370	AX298370 Sequence
26	217	10.8	3489	6 AX239647	AX239647 Sequence
27	190.4	9.5	10578	6 AB065618	AB065618 Homo sapi
28	190.4	9.5	102539	9 AL591866	AL591866 Human DNA
29	190.4	9.5	191481	2 AC108123	AC108123 Homo sapi
30	187	9.3	4116	5 AF406649	AF406649 Squalus a
31	187	9.3	4134	6 AR182343	AR182343 Sequence
32	187	9.3	4134	6 AX417770	AX417770 Sequence
33	187	9.3	4134	6 AX449101	AX449101 Sequence
34	185.4	9.2	1383	9 AF387619	AF387619 Homo sapi
35	184.2	9.2	1101	6 AX282896	AX282896 Sequence
36	184.2	9.2	93897	2 AC129831	AC129831 Rattus no
37	182.8	9.1	210	9 AF458149S4	AF458149S4 Homo sapi
38	180.2	9.0	2729	6 AX354026	AX354026 Sequence
39	172.6	8.6	5249	4 S67307	S67307 Ca(2+)-sens
40	172.6	8.6	5275	6 AR012622	AR012622 Sequence
41	172.6	8.6	5275	6 AR028465	AR028465 Sequence
42	172.6	8.6	5275	6 AR078217	AR078217 Sequence
43	172.6	8.6	5275	6 AR177781	AR177781 Sequence
44	172.6	8.6	5275	6 I75051	I75051 Sequence 1
45	168.6	8.4	173724	2 AC026766	AC026766 Homo sapi

ALIGNMENTS

RESULT 1

LOCUS AR208617 2010 bp DNA linear PAT 20-JUN-2002

DEFINITION Sequence 8 from patent US 6383778.

ACCESSION AR208617

VERSION AR208617.1 GI:21509818

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 2010)

AUTHORS Zuker,C.S., Adler,J.E. and Lindemeier,J

TITLE Nucleic acids encoding a G-protein coupled receptor involved in sensory transduction

JOURNAL Patent: US 6383778-A 8 07-MAY-2002;

Pred. No. is the number of results predicted by chance to have a

```

FEATURES { Location/Qualifiers
    source             1..2010
                        /organism="unknown"
BASE COUNT      387 a      691 c      515 g      417 t
ORIGIN
Query Match      100.0%; Score 2010; DB 6; Length 2010;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2010; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCACTACAGCCGATGAGATGAGTGGAGACAGAGGAGGCTTCCGGCTTGGTCTG 60
DB 1 ATCACTACAGCCGATGAGATGAGTGGAGACAGAGGAGGCTTCCGGCTTGGTCTG 60
QY 61 CGTACACACCCAGCCGCGACACACAGTGGAGCCATGGTGCAGTGTATGTCACATTC 120
DB 61 CGTACACACCCAGCCGCGACACACAGTGGAGCCATGGTGCAGTGTATGTCACATTC 120
QY 121 CGCTGGAACTGATGATGTTGCTGTGTGAGACAGACACTATGGCCGCGACATGGCCAG 180
DB 121 CGCTGGAACTGATGATGTTGCTGTGTGAGACAGACACTATGGCCGCGACATGGCCAG 180
QY 181 CTGCTTGGCAGAGCGCGCGCGCGCGAGATGATGATGCTTCCAGAGAGCGTGGCC 240
DB 181 CTGCTTGGCAGAGCGCGCGCGCGCGAGATGATGATGCTTCCAGAGAGCGTGGCC 240
QY 241 ACACTGACGCCACCAACCAACATGATGATGAGAGAGCGCCAGCGCTGTGACCATTTGTG 300
DB 241 ACACTGACGCCACCAACCAACATGATGATGAGAGAGCGCCAGCGCTGTGACCATTTGTG 300
QY 301 GACAAGTGGAGAGACAGACAGAGCGCGTGTGTGTGTTCTGCGCCGACCTGACCCCTG 360
DB 301 GACAAGTGGAGAGACAGACAGAGCGCGTGTGTGTGTTCTGCGCCGACCTGACCCCTG 360
QY 361 TACCACTTCTTCAATGAGTGTGCTGCGCAGAACTTCAAGGAGCGCGTGGATGGCTCC 420
DB 361 TACCACTTCTTCAATGAGTGTGCTGCGCAGAACTTCAAGGAGCGCGTGGATGGCTCC 420
QY 421 GAGTCTGGGCGATCGACCGCGTCTGACAACTCAAGAGTGGGCACTTGGGACCC 480
DB 421 GAGTCTGGGCGATCGACCGCGTCTGACAACTCAAGAGTGGGCACTTGGGACCC 480
QY 481 TTCTGGGCAATCCATCCAGAGGTCCTATCCCGGGCTTCAAGTGTCCCGGATGG 540
DB 481 TTCTGGGCAATCCATCCAGAGGTCCTATCCCGGGCTTCAAGTGTCCCGGATGG 540
QY 541 GGGCCACAGGCTGGGCGCGCCACCCCTCAGACAGACAGACGATATACCTGCAACAG 600
DB 541 GGGCCACAGGCTGGGCGCGCCACCCCTCAGACAGACAGACGATATACCTGCAACAG 600
QY 601 GAGTGGCAACAGTGCCTGGAACGCCACTTCTCTTCAACACCATTTGAGGCTCTGGG 660
DB 601 GAGTGGCAACAGTGCCTGGAACGCCACTTCTCTTCAACACCATTTGAGGCTCTGGG 660
QY 661 GAGGCTGTGCTACAGCGTACTCTGCGGTATGCTGTGGCCATGGCCCTGCAACAGC 720
DB 661 GAGGCTGTGCTACAGCGTACTCTGCGGTATGCTGTGGCCATGGCCCTGCAACAGC 720
QY 721 CTCTCGGCTGTGCAAAACACCTGACCAAGAGGTTGTTCTTACCCCTGGACACTGTT 780
DB 721 CTCTCGGCTGTGCAAAACACCTGACCAAGAGGTTGTTCTTACCCCTGGACACTGTT 780
QY 781 GAGAGATCTGGAAGGTCAACTTCACTCTCTGAGACCAACAAATCTTCTGACCCGCAA 840
DB 781 GAGAGATCTGGAAGGTCAACTTCACTCTCTGAGACCAACAAATCTTCTGACCCGCAA 840
QY 841 GGGAGCTGGCTCTGCACTTGGAGATTTGTCAATGGGAATGGGAGCCGAGACCAAAATCCC 900
DB 841 GGGAGCTGGCTCTGCACTTGGAGATTTGTCAATGGGAATGGGAGCCGAGACCAAAATCCC 900
QY 901 TTTCAGAGCTGCGCTCTCTACTACCCCTGACAGACAGTGAAGACATCAAGACTCT 960
DB 901 TTTCAGAGCTGCGCTCTCTACTACCCCTGACAGACAGTGAAGACATCAAGACTCT 960

```

```

QY 961 CTGACACCGCTCAACAACAGATCCATATGTCATGTTCAAGAGGTCCAGTCAGGG 1020
DB 961 CTGACACCGCTCAACAACAGATCCATATGTCATGTTCAAGAGGTCCAGTCAGGG 1020
QY 1021 CAAAAGAAAGAGCTGTGGGATTCACAGTGTGCTGCTGAGTGTACATGCTTCC 1080
DB 1021 CAAAAGAAAGAGCTGTGGGATTCACAGTGTGCTGCTGAGTGTACATGCTTCC 1080
QY 1081 GGCACCTTCTCAACCACTGATGATGCGGATTAACAGAGTGTCTCCAGAGTGAACC 1140
DB 1081 GGCACCTTCTCAACCACTGATGATGCGGATTAACAGAGTGTCTCCAGAGTGAACC 1140
QY 1141 TCGTCTTCAAGCGGACAGTGTCTTCTGTAATGGCATGAGGACACCATGCTGTG 1200
DB 1141 TCGTCTTCAAGCGGACAGTGTCTTCTGTAATGGCATGAGGACACCATGCTGTG 1200
QY 1201 GCCCTGCTGGCCGCTTCCGAGTGTCTGACACACCTGGGCACTCTGTGATTTCTGAGG 1260
DB 1201 GCCCTGCTGGCCGCTTCCGAGTGTCTGACACACCTGGGCACTCTGTGATTTCTGAGG 1260
QY 1261 CACTTCCAGACACCCATAGTTGCTGGGTGGGGGCCCATGTGCTTCTGATGCTGACA 1320
DB 1261 CACTTCCAGACACCCATAGTTGCTGGGTGGGGGCCCATGTGCTTCTGATGCTGACA 1320
QY 1321 CTGCTGTGGTGGATACATGATGTTGCCGCTGACGTGAGCGGGCGCCCAAGTCTCCACC 1380
DB 1321 CTGCTGTGGTGGATACATGATGTTGCCGCTGACGTGAGCGGGCGCCCAAGTCTCCACC 1380
QY 1381 TGCTCTGCGCGCCAGAGCGCTTCCCTGCTGCTGCAATTTGCAATCTCTGTATCGCC 1440
DB 1381 TGCTCTGCGCGCCAGAGCGCTTCCCTGCTGCTGCAATTTGCAATCTCTGTATCGCC 1440
QY 1441 GTGCGTCTTCCAGATGCTGCGGCTTCAAGATGAGGAGCGGCTTCCACGCGCTAC 1500
DB 1441 GTGCGTCTTCCAGATGCTGCGGCTTCAAGATGAGGAGCGGCTTCCACGCGCTAC 1500
QY 1501 AGCTACTGGGTCGCTACACAGGGGCCCTTACGTCTATATGSCATTTATACGGTACTCAA 1560
DB 1501 AGCTACTGGGTCGCTACACAGGGGCCCTTACGTCTATATGSCATTTATACGGTACTCAA 1560
QY 1561 ATGGTCAATTTGGAATTTGATGATGCTGGACAGGGGCTGACGCCACCGCTACTGACGCC 1620
DB 1561 ATGGTCAATTTGGAATTTGATGATGCTGGACAGGGGCTGACGCCACCGCTACTGACGCC 1620
QY 1621 GATGACCCCAAGATCAACAATTTCTCTGTAAACCCAACTACCGCAACAGCTGTGTTTC 1680
DB 1621 GATGACCCCAAGATCAACAATTTCTCTGTAAACCCAACTACCGCAACAGCTGTGTTTC 1680
QY 1681 AACACAGGCTGGAAGTGTGCTGTCTAGTGTGGGTTTCAAGCTTCCGCTACATGGGCAAA 1740
DB 1681 AACACAGGCTGGAAGTGTGCTGTCTAGTGTGGGTTTCAAGCTTCCGCTACATGGGCAAA 1740
QY 1741 GAGTGGCCCAACAATCAACAGAGGAGGCAAGTATATACCCCTGACAGTACTTATTTCC 1800
DB 1741 GAGTGGCCCAACAATCAACAGAGGAGGCAAGTATATACCCCTGACAGTACTTATTTCC 1800
QY 1801 ACCTCATCCGCTTCCCTGCTGACCTTATGCTGTGCTTACAGCGGGGCTGTGCTACCATC 1860
DB 1801 ACCTCATCCGCTTCCCTGCTGACCTTATGCTGTGCTTACAGCGGGGCTGTGCTACCATC 1860
QY 1861 GTGAGCTCTTGGTCACTGTGCTAACCTCTGCGCATCAAGCTGGGCTACTTCTGAGGCCCC 1920
DB 1861 GTGAGCTCTTGGTCACTGTGCTAACCTCTGCGCATCAAGCTGGGCTACTTCTGAGGCCCC 1920
QY 1921 AAGTGTACATGATGATCTCTTCACTACCGGAGGAGCAACAGCGCGCTACTTCAACAGCATG 1980
DB 1921 AAGTGTACATGATGATCTCTTCACTACCGGAGGAGCAACAGCGCGCTACTTCAACAGCATG 1980
QY 1981 ATCCAGGCGTACACATGAGAGGAGACTAG 2010
DB 1981 ATCCAGGCGTACACATGAGAGGAGACTAG 2010

```

RESULT 2
LOCUS AR208615 2993 bp DNA linear PAT 20-JUN-2002
DEFINITION Sequence 3 from patent US 6383778.
ACCESSION AR208615
VERSION AR208615.1 GI:21509815
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2993)
AUTHORS Zuker,C.S., Adler,J.E. and Lindemeyer,J.
TITLE Nucleic acids encoding a G-protein coupled receptor involved in
sensory transduction
JOURNAL Patent: US 6383778-A 3 07-MAY-2002;
FEATURES
source location/Qualifiers
1..2993
BASE COUNT 619 a 962 c 729 g 683 t
ORIGIN
Query Match 61.3%; Score 1232.8; DB 6; Length 2993;
Best Local Similarity 76.9%; Pred. No. 9.6e-194;
Matches 1568; Conservative 0; Mismatches 442; Indels 30; Gaps 4;
QY 1 ATCACTACAGGCGCATCAGCATAGTGGAGACAAGGTGCGCTCCCGCTTGGTG 60
DB 505 ATCACTACAGGCGCATCAGCATAGTGGAGACAAGGTGCGCTCCCGCTTGGTG 564
QY 61 CGTACCAACACCGGCGGACGACGACGACGACGACGACGACGACGACGACGACG 120
DB 555 CGCAAGTGGCCAGGCGGACGACGACGACGACGACGACGACGACGACGACGACG 624
QY 121 CGCTGAACTGGATCATCTGCTGTGAGACGACGACGACGACGACGACGACGACG 180
DB 625 CAATGAACTGGATCATCTGCTGTGAGACGACGACGACGACGACGACGACGACGAC 684
QY 181 CTGCTTGGCGAGCGGCGTGGCC---CGGCGGACATCTGCTGCTGCTGCTGCTG 237
DB 665 CTGTTGACCGACGCTTCAACAAAGAGCGACATCTGCTGCTGCTGCTGCTGCTG 744
QY 238 CCCACATGACGCGGACGACGACGACGACGACGACGACGACGACGACGACGACG 297
DB 745 CCCATACGAGTCCAGGCGGACGACGACGACGACGACGACGACGACGACGACGACG 804
QY 298 GTGACAAAGTGGACGACGACGACGACGACGACGACGACGACGACGACGACG 357
DB 805 CTGGACAAAGTGGACGACGACGACGACGACGACGACGACGACGACGACGACG 864
QY 358 CTGTACCACTTCTCAATGAGGTGCTGGCCAGACTTCACGCGGCGCGTGTGATCCG 417
DB 865 CTGTACCACTTCTCAATGAGGTGCTGGCCAGACTTCACGCGGCGCGTGTGATCCG 924
QY 418 TCCGAGTCTGGGCGGACGACGCGGCTGCTGACAACTTCACGAGAGCTGGGCGCTT 477
DB 925 TCTGAGTCTGGGCGGACGACGCGGCTGCTGACAACTTCACGAGAGCTGGGCGCTT 984
QY 478 ACCTTCTGGGCGGACGACGAGGCTGCTGACGAGGCTGCTGAGTGTGCGGAG 537
DB 985 ACTTCTTGGGCGGACGACGAGGCTGCTGACGAGGCTGCTGAGTGTGCGGAG 1044
QY 538 TGGGCGGCGGAGGCTGGGCGGCGGACGCGGCGGACGAGGCGGAGGCTGCGGAG 597
DB 1045 CGCGGTGACAAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1104
QY 598 CAGGAGTGGAGACTGCTGAGGCGGACGCTGCTTCAACACCATCTTCAAGGCTCT 657
DB 1105 CAGGAGTGGAGACTGCTGAGGCGGACGCTTCAACACCATCTTCAAGGCTCT 1164
QY 658 GGGGAGGCTGTGCTTACAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
DB 1165 GGGGAGGCTGTGCTTACAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1224

QY 718 AGCTCTCGGCTGTGACAAAAGACCTGACACGAGGCTGTGCTTACCCCTGGACGCTG 777
DB 1225 AGCTCTCGGCTGTGACAAAAGGCTGCGGTGACGACAAAGGCTGTGCTTACCCCTGGACGCTA 1284
QY 778 CTGAGGAGATCTGAGAGTCACTTCACTGCTGAGACGACCAATCTTCTTTCAGCCG 837
DB 1285 CTCAGGAGATCTGAGAGTCACTTCACTGCTGAGAGGCTGCTTCTTTCAGCCG 1344
QY 838 CAAGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 897
DB 1345 CAAGGGAGATCTGAGAGTCACTTCACTGCTGAGAGGCTGCTGCTGCTGCTGCTGCTGCT 1404
QY 898 CCTTTCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 954
DB 1405 CCTTTCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1464
QY 955 ACATCTGACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1014
DB 1465 GTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1524
QY 1015 TCAGGGCAAAAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1074
DB 1525 CCAGGGCAAAAGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1584
QY 1075 CTTCGGGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1113
DB 1585 ATGCGGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1644
QY 1114 AACGAGTGTCTTACAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1173
DB 1645 TCCATGTGTCTTCAAG 1704
QY 1174 TGGCATGAGGACCAACAG 1233
DB 1705 TGGCAGAGAGTGGCCAG 1764
QY 1234 CTGGGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1293
DB 1765 CTGGGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1824
QY 1294 GGCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1353
DB 1825 GGCCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1884
QY 1354 TACGTGGGCGGCGGACGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1413
DB 1885 TATGTGGGCGGCGGACGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1944
QY 1414 TTCACAATTTGCAATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1473
DB 1945 TTCTCATCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2004
QY 1474 ATGGCAGCGGCTTCCGACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1533
DB 2005 ATGGCAGCGGCTTCCGACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2064
QY 1534 TCTATGCGATTTATACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1593
DB 2065 TTTGCGGCTTTCATACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2124
QY 1594 CCGTACGCGGACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1650
DB 2125 ACCATCAACCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2184
QY 1651 AACCCCACTACCGGACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1710
DB 2185 CACCTCACTACCGGACGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2244
QY 1711 GTGGGTTTACGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1770
DB 2245 CTGGGTTTACGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2304

[illegible]

Query Match	Similarity	61.2%	Score 1230:	DB 103:	Length 2529:
Best Local	Similarity	76.9%	Prod. No. 2.8e-193:		
Matches 1564:	Conservative	0:	Mismatches 440:	Indels 30:	Gaps
QY	1	ATCACCTACAGCGCCATCATGCGATGAGCTGTGACAGACAAAGGTGCGCTTCCGGCTTTGCTG	60		
Db	493	ATCACATATACAGCGCCATCTCCGACAAAGCTGGGGAGCAAGCGGACACTTCCCTGACATGCTA	552		
QY	61	CGTACACACCCAGAGGCCGACACGACGCTGAGGGCCATGATGCTGATGCTGACCTTC	120		
Db	553	CGCACAGTGTCCAGGGCCACCCACACATCTGAGGGCCATGCTGATGCTGATGCTGACCTTC	612		
QY	121	CGCTGGAATCTGATCTATGTGTGCTGTGACGACGACACTATGCGCGGACACATGGCCAG	180		
Db	613	CAATGGAATCTGATTTGT	672		
QY	181	CTGCTTTGGGAGACGCGTGGCC---CGGGCGACATCTGTGATGCGCTTCCAGGAGACGTG	237		
Db	673	CTGTGTGAGCGACGCTGTGTGACAAAGACGCGACATCTGTGATTTGCTTCCAGGAGGTTTCT	732		
QY	238	CCACACTCTGACCCCAACGACAAATGAGCTGAGAGGAGCGGCGACGCGCTGTGTGACATTT	297		
Db	733	CCCATATCTGATGATCCAGCGGCTCATGAGAGTGTGAGAGACAGACAACTGTGACAAATTC	792		
QY	298	GTGACAGAGCTGTGACAGACAGACAGCGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	357		
Db	793	CTGGAGCAAGCTCTGGGGGACCTTGGGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	852		
QY	358	CTGTACCAATCTTCTCAATAGGTGTGCTGCGCGCCAGAAATTCAGGGGCGCCCTGTGTGATCG	417		
Db	853	CTGTATAGTCTTTTTCACGAGGT	912		
QY	418	TCCGAGTCTGTGGGACATGACCCGGTCTGTGACAACTTCAGGAGCTGGGCACTTTGGGCT	477		
Db	913	TCTGTAGTCTGTGGGCTATCCAGCCAGTTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	972		
QY	478	ACCTTCCTGTGGGATACCAATTCGAGGCGGCGCCATCCGGGGCTGACGTAGTGTGCGCGAG	537		
Db	973	ACTTTTCTGTGGGCTACCAATTCGAGGCGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1032		
QY	538	TGGGGCCACAGGCTGTGGGGCGCCACCCCTCAGACGAGGACGCGACAGCTATACCTGTGCAC	597		
Db	1033	CGCCGTGACAAACGACAGGGGATATCCCGGTGCTTAACAGACCACTGTGGGAGCACTGTGCAC	1092		
QY	598	CAGGAGTGTGACAACTGTGCTGTGAAGCCACACTTGTCTTAAACATTTCTGAGGCTCTT	657		
Db	1093	CAGGAGTGTGACGCTGTGCTGTGAAGCCACCAAGTCTTCAACAACTTCTTATCTTGTG	1152		
QY	658	GGGAGAGCTGTGTGTACAGCTGT	717		
Db	1153	GGGAGAGCGGT	1212		
QY	718	AGCTTCCTGT	777		
Db	1213	AGACTTCCTGT	1272		
QY	778	CTTGTGAGGATGTGTGAGGCTCAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	837		
Db	1273	CTTGTGAGGATGTGTGAGGCTCAACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	1332		
QY	838	CAGGAGGAGCTGT	897		
Db	1333	CAGGAGGAGCTGT	1392		
QY	898	CCCTTCACAGGCTGT	954		
Db	1393	CCCTTCACAGGCTGT	1452		
QY	954	ACATCTGT	1014		
Db	1453	GT	1512		
QY	1015	TCAGGAGCAAAAGAGGCTGT	1074		

Db	1513	CGAGGCGAAATGAAAAGTCTTGAGGCTTCACCCCTTGTTGCTTGAGAGTCTTGATATGT	15172
Oy	1075	CTTCCCGGCACTCTCTCCCAACCACTAGAA-----TGCCGAAAT	11133
Db	1573	ATGCGAGGCAACGTACCTCAACCGCTCAGAGATGATTAAGTTAACTGTCTGTCTGCCCGGGT	16322
Oy	1114	AACGAGTGTCTTCCACGAGTGAAGACTCTGCTTCAAGCGCAGCTGTCTTCTTGGA	11733
Db	1633	TCGATGTGTCTTCAAGAAAGCAATCACTGTCTTCCACGGCGGCCTACCTCTCTGAG	16922
Oy	1174	TGGATATAGGCAACCACTATCGCTGTGGCCCTCTGGCGCCCTGGGCTTCTTGACAC	12333
Db	1693	TGGCAGGAATGCCACACATCTGTGTGGCCATTAAGTGTGCTTGCCCTGTCTTCAAGTAA	17522
Oy	1234	CTGGCATCTCTGGTATATTTCTGAGGCACTTCCAGACACCAATAGTTGCTGGGCTGG	12933
Db	1753	CTGGCATCTCTTTCATCTTCTGAGACATTTCCAAACCACTGATGGTGGCTGGCGCGGT	18122
Oy	1294	GGCCCCATGTGCTCTCGATGCTGACACAGCTGTGCTGGTGGCAATAGTGTGTCCCGGT	13533
Db	1813	GGCCCCATGTGCTCTCGATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG	18722
Oy	1354	TAGCTGGGGCGCGCCAGAGTCTCACCTGCTGCTGTGCGCGAGCGCCCTCTTTCCTCTGC	14133
Db	1873	TATGTGGGGCGCGCCACAGGCTCTCTATGCTCTTGTGCGACAGGCTTCTCACGCTGTGC	19322
Oy	1414	TTTCATATTTGCAATTCCTGTATGCGCGGTGTCTTTCAGATGCTGTGCGGCTTCAAG	14733
Db	1933	TTTCCATCTGTCTATCTGTCAATCACCGGTCTCTTCCAGATGTGTGTCTTCAAG	19922
Oy	1474	ATGCGCAGCGCTCTCCACCGCGCTCACTGAGTGTGGTCCGTACAGAGGGCGCTCACTGC	15333
Db	1993	ATGCGCAGAGCGCTCGCCAAAGTGTGCTTATGATGCTGTACACAGGGCGCTATGTC	20522
Oy	1534	TCATATGCAATTTATCAACAGTACTCAAAATGTGTATGTGTATGTGCAATGTGGCAGG	15933
Db	2053	TTGTGCGCTTCAACAGGCAAGGTGTGCGCTGTGTGTGTGGCAACATGTGTGGCCACC	21122
Oy	1594	CCCTAGTCCCAACC-----CGCTACTGACCCCGATGACCCCAATGACAAATGTCTCTCGT	16502
Db	2113	ACGATCAACCCGANTGGCGCGAGACCGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	21722
Oy	1651	AAACCCCACTACGCAACAGCTGTGTGTCAACACCAAGCTGTGAGCTGTGTCTCAAGT	17102
Db	2173	CACCTTACTACGCAACAGGCTGTGTGTCAACACCAAGCATGATGTGTGTGTGTGTGTGTGT	22322
Oy	1711	GTTGGTTTCAAGTCTGCTCAATGGGCAAGAGCTGCCCAACATCAACAGAGGCGAAG	17702
Db	2233	CTGGGTTTCAAGCTTCTGTGTCAATGGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	22922
Oy	1771	TTTCATCAACCTCAAGATGACCTCTTATTTACCTCATTCGCTCTCTGTGACCTTATG	18302
Db	2293	TTTCATCACTGTCAAGATGACCT	23522
Oy	1831	TTCTGTCAACAGCGGT	18902
Db	2353	TTCTGTCAACAGCGGT	24122
Oy	1891	CTGGCATCAAGCTGT	19502
Db	2413	CTGGCATCAAGCTGT	24722
Oy	1951	CGCAACACGCGGCTTACTTCAACAGATGATTCAGGGCTACACATGAGAGAG	2004
Db	2473	CGCAACACGCTACGCTATTTCAATACATGATTCAGGGCTACACATGAGAGAG	2526

Accession	Gene	Length	Feature	Source
AF337041	<i>Mus musculus</i> candidate taste receptor T1R2 cds.	3060 bp	mRNA	ROD 25-APR-2001
LOC037041	<i>Mus musculus</i> taste receptor T1R2	linear	mRNA, complete	

ACCESSION	AF337041
VERSION	AF337041.1
KEYWORDS	GI:13785658
SOURCE	Mus musculus.
ORGANISM	Mus musculus. Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 3060) Montmayeur,J.P., Liberles,S.D., Matsumami,H. and Buck,L.B. A candidate taste receptor gene near a sweet taste locus Nat. Neurosci. 4 (5), 492-498 (2001)
AUTHORS	21219400
JOURNAL	11319557
PUBLISHED	2 (bases 1 to 3060) Montmayeur,J.-P., Liberles,S.D., Matsumami,H. and Buck,L.B. Submitted (18-JUN-2001) Neurobiology, Harvard Medical School, 220 Longwood Avenue, Boston, MA 02115, USA
REFERENCE	Location/Qualifiers 1..3060 /organism="Mus musculus" /strain="C57BL/6J" /db_xref="taxon:10090" /chromosome="4" 1..3060 /gene="Tlr2" 32..2563 /note="G protein coupled receptor" /codon_start=1 /product="candidate taste receptor Tlr2" /protein_id="AAK39438.1" /db_xref="gi:13785659"
TITLE	/translation="MGPOARTILFLFLIALALPKPVMLVGNSDPHLAGDYLIGLFTL HANKYNSHLSTLOVPKCNEYNMKLVLSNYIMQARFAVEIINNCSSILRVLGLGYEMV DYCLISNNIDPGTLFSLQIDDFLLDKSOTRQOVAIVGPDSBAITVSNLSTF LVPQVTSATIDKLRLDRREPPEMLRPVSATIHLEAVOLVHEQVMWIVLVSDDY GRESHSLSORLTNTGDICIAFEVLVPEPNQAREDDODLIDKLRTSARV VIPELSLNHFREVLNMTGFVWLASBSMAIDPYLNLTELRLHTGTFLGYTORV SIPEOSFRVHKDEYPMPIETSLRTCCODCAKNITSEFNVMILSGERYYSV YSAYVAHLIRHLHCNOVCSTOIYYPMQLREIHWNVFTLLGNLFEPDGDMAM LDLIIONOMGISQNPFOSIASYSTETRLTYISNVATTPTNPVTIPISMGRSCOPGM KRPLGVLPACECEVDCCPPGYILNRSVDEFNCLSPGSMSWSYKNNAICFKRRALFELWH EVPIPIVITLAALGISTLAILLFMHPQTPIPWRSAGGPCILMAPLIAAGMPV YVGPVFESCRCROAFPVTFESVSLCIYSRFQIVCFKMARLPSAYGFMMKHBP YVFAPIITAVVALVAGNMALATTINPIGRDDPDPIIIISCHPYRNGLIFPTSMDL LSTVGESFAVKEGLEPTNNNEAFITLSMTFESTSSISICTPSYSDVLYTIIMDLL VTVNLAIIGTGIFPCYMIILFPERTSVATFNSMTGITYMRKS"
BASE COUNT	613 a 990 c 751 g 706 t
ORIGIN	
Query Match	60.7%; Score 1220.6; DB 10; Length 3060;
Best Local Similarity	76.5%; Pred. No. 9.8e-192;
Matches 1560; Conservative	0; Mismatches 449; Indels 30; Gaps 4;
2	TCACCTACAGCGCCATCAGCATGAGTGCAGAGACAAGTGCAGTCCCGCGCTTTGGTCG 61
D5	525 TCACATATAGGCCCCATCACCGCAGAAGCTGCGAGAACAGCGCGCTTCCCTGCATGCTGC 584
OY	62 GTACCACACCGAGCGCGAGACCAACGATGAGAGGCGCATGAGTGCAGTTCGCATTCG 121
D5	585 GCACGTGCCCCAGCGCCACCAACCAACATCGAGGCGCATGTGCTCACTGATGCTTCC 644
OY	122 GCTGAAGTGCATATTTGCTGCTGTGAGACAGCAGCAACTATAGGCCCGAGCAATAGGCCAGC 181
D5	645 AATTGAAGTGTGCTGCTGTGCTGTGAGACGATGAGATTAATGCGGAGAGAAACAGCACCC 704
OY	182 TGCTGGGAGAGCGCGTGC-----CGGCGCGCAATCTGATCGCCTCCAGAGAGCGTGC 238
D5	705 TGCTAGGACGAGCTGTGACCAACACTGCGCATATCTGATTTGCCCTTCCAGAGAGGTTCTGC 764
OY	239 CCACACTCGAGCCCAACGAGAATGATGATGAGAGAGAGCGCCAGCGCTTGCTGACATTTG 298

D	b	765	CTGTACCGAAGAACCCACAGGCGGTGAGGCGCTGAGGAGACAGGACCAACTGACAAACATCC	824
Q	y	299	TGGACAACTCTACACAGACACAGCGCGCGTGTGCTGTCTCTGCGCCAGCTGACCC	358
D	b	835	TGGAAACCTCTCGCGGAGACCTCGCGCGTGTGTGTATCTTCGCGACAGAGCTGAGCC	884
Q	y	359	TGTACCACTTCTTAAATGAGGTGTGGCGGAAATTCACGGGCGCGGTGTGGATCGGCT	418
D	b	885	TGCACAACTTCTTCCGGAGGTGTGTGGCTGAGAACTTCACAGGCTTGTGTGGATTCCT	944
Q	y	419	CCGAGTCTCTGGGCCCATGAGCCCGGCTCTGTGCACAACTCCAGAGACTGAGCGCCACTTGGCA	478
D	b	945	CTGAGTCTCTGGGCCCATGAGCCCTTCTTACAACTCCACAGAGCTGCGCCACAGGGCA	1004
Q	y	479	CCCTTCTCTGGGCATACCATCCAGAGCGTCCCATCCCGGGCTTCAGTAGTTCGCGAGT	538
D	b	1005	CTTTCCTCTGGGCTACCATCATCCAGAGGCTGTCCATCCCTTGCTCAGCCAGTTCGAGTGC	1064
Q	y	539	GGGGGCCCAAGGCTGGGCGCCGACCCCTCAGACAGAACACAGAGACTTACTTCGAAC	598
D	b	1065	GCCACGACAAACCCAGAGATTCACATGCTATGACAGACACAGCCCTGAGAGCTACTTAAAC	1124
Q	y	599	AGGATGTGGACAATCTGCCTGAGCCACCTTGTCTTCAACACATCTCAGAGCTCTGCG	658
D	b	1125	AGGATGTGAGCGCTCGATGAACTACACAGATCTCTTAAACAGCTTCTCAATGCTTTGG	1188
Q	y	659	GGGAGCGTGTCTATACAGCGCTGTACTCTGCGGTATAGCTGTGGCCATGCGCTGCACA	718
D	b	1185	GGGAGCGTGTGTCTATACAGTGTATCTAGCGCGCTCTACGGGGTATGCCACACTTCACAA	1244
Q	y	719	GGCTCTCTGGGTGTGACAAAAGCACCCGACACAAAGAGGTGTACCTCCCTGGCAGCTGC	778
D	b	1245	GACTCTCTCACTGACACCAAGGTCTCGCTGCACAAAGAAATCGTATCATCATGCGAGCTAC	1304
Q	y	779	TTGAGGAGATCTGGAAGTCAACTTCACTCTCTGGACCAACAACTCTTCTCGACCGC	838
D	b	1305	TCAGGAGATCTGGCAGATGTCAACTTCACGGCTCTCGGGCAACACAGCTCTTCTTCGAGAAC	1364
Q	y	839	AAGGGAGCTGTGGCTCTGACTTGTGAGATTTCTCAGTGTGCATATGGACCGGACCGAATTC	898
D	b	1365	AAGGGAGATCTCCCATCTCTCTGTGACATCTCAGTGTGCATATGGGGCTCGACCGAAGAC	1424
Q	y	899	CCCTTCACAGAGGTGCGCTCCACTACCCCTGAGGAGCAAGCTGGAAGAAACAT---CAAGA	955
D	b	1425	CCCTTCACAGAGATGCGCTCTACTCTCCCAACGAGACAGAGCTGACCTTACATTAAGCAATG	1484
Q	y	956	CATCTCTGCACCCGTCACAAACACAGATCTCTATGTCTCATATGTTTCCAAAGAGTCCAGT	1015
D	b	1485	TGTCTGTGTACACCCCAACACAGAGTCTCCCATATCATATGTTTCTTAAGATTTCCAGAC	1544
Q	y	1016	CAGGGCAAAAGAGAGAGCGTGTGGGCACTCCAGCTGTGCTTCGATGCATCGACTGCC	1075
D	b	1545	CTGGGCAAAATGAAAAAACCATATAGGCTCTCCACCCGTGCTCGATGTGTGAGACTGTC	1604
Q	y	1076	TTCCCGGCACTTCTCTCAACCAACACAGAA-----TGCCTCGAATA	11144
D	b	1605	CGCGGGGCACTTACTCTCAACCGATCAGTAATAGTTTAAGTCTGTCTCTGCCCCGGGTT	1664
Q	y	1115	ACGATGTCTCTTACACAGTGAAGACCTCGCTTCAAGGGGAGCGGTCTTCGAGAT	11747
D	b	1665	CCATGTGTCTTACAAACAAACATGTGCTTCTTCAAGGGGGGCTGGGCTCTCTGTGAGT	1724
Q	y	1175	GGCATGAGGACCCACCATCGCTGTGGCCCTGTGTGGCCGCCCTGGGGCTCTCTCACACCC	12344
D	b	1725	GGCAGCAAGTCCCACTATGTGTGTGACATCTCTGGCGCCCTCGGGCTTCATCACTAATAGC	17844
Q	y	1235	TGGCCATCTCTGTGTATATTCTGAGAGCACTTCAGACACCATATAGTTGCTCTGGCTGGGG	12944
D	b	1785	TGGCCATCTCTGTCTCATCTTCTGTGGAGCAATTTCCAGACGCCCATGTGGCGCTCTGGCGGCG	18444
Q	y	1295	GGCCCATGTGCTTCTCATATCTCAGACACTGTGTGTGTGGCATATAGTGTGTGCGGTGT	13544
D	b	1845	GGCCCATGTGCTTCTCATATCTCTGTGTGCGCTCTGTGTGTGGATGTTCGGATGTGTGCCGTGT	19044

QY	1355	ACGTGGGGGGGGCCAAAGGCTCTACCTGGCTGGGGCCAGGAGCCCTTCCCTGGT	1414
QY	1355	ACGTGGGGGGGGCCAAAGGCTCTACCTGGCTGGGGCCAGGAGCCCTTCCCTGGT	1414
Db	1905	ATGTGGGGCCCCCAGGGCTTCTCTCTGTCTTGTCTGGCCAGGCTTCTACCGCTTGGT	1964
QY	1415	TCACATTTTGCATCTCCCTGTATCCGCGTGGCTCTTCTTCCAGATTCGTCTGGCTTCAGA	1474
Db	1965	TCCTCGGTCTGCTCTCTCTACATCAGGTGCGCTCTCTCCATTCAGATTTGTGTCTTCAGAA	2024
QY	1475	TGGCCAGCCGCTTCCACGCGCTTACAGTACTGGTCCGCTACCAAGGGGCGCTTACGTCT	1534
Db	2025	TGGCCAGAGCGCTGGCCAAACGCGCTACGGCTTCTGTGATGTGGTACCAAGGCGCGCTACGTCT	2084
QY	1535	CTATGGCATTTTATTCACGGGACTCAAAATGGCATTTGGTATTTGGCAATGCTGGACGCG	1594
Db	2085	TGTGTGGCTTCTATTCAGAGCGGCTTCAGAGTGGCCCTGTGGCAGCAACATGCTGGCCACCA	2144
QY	1595	CTCAGTCCCAACC---CCGTACTGACCCCGATGAGCCCAAGATTCACAAATTTCTCTCTGTA	1651
Db	2145	CCATCAACCCCATTTGGCCGGGAGCCGAGCCCGATGAGCCCAATTCATTAATCTCTCTGGCC	2204
QY	1652	ACCCCACTACCGCAACAGCTGCTGTTCACACACAGCGCTGACCTGCTCTCTACGTGG	1711
Db	2205	ACCCCACTACCGCAACAGCGGCTACTCTTCACACAGCAGCTTGGCTGCTGCTGCTGC	2264
QY	1712	TGGGTTTTCAGCTTGGCGCTTACATGGGCAAAAGACTGGCCCAACTACAGACGAGGCGCAAGT	1771
Db	2265	TGGGTTTTCAGCTTGGCGCTTACATGGGCAAAAGACTGGCCCAACTACAGACGAGGCGCAAGT	2324
QY	1772	TCATCAACCTCAGATGAGACTTCTTATTCACCTCATCCGCTCTCCCTGACACTTCATGT	1831
Db	2325	TCATCAACCTCAGATGAGACTTCTTCTCTTCACTCTCCATCTCCCTCTCAGACTTCATGT	2384
QY	1832	CTGCTTACAGGGGGGCTGCTGTCACATCGTGGACTCTTGGTCTCATCTGCTCAACCTCC	1891
Db	2385	CTGCTCAGATGGCGCTGCTGTCACATCGTGGACTCTTGGTCTCATCTGCTCAACCTTC	2444
QY	1892	TGGCCATCAGCGCTGGGCTTACTTGGGCCCCCAAGTGGTACATGATTCCTTCTACCCGAGC	1951
Db	2445	TGGCCATGCGCTTGGGGTACTTGGGCCCCCAAGTGGTACATGATTCCTTCTACCCGAGC	2504
QY	1952	GCAACAGCGCGCGCTTACTTCAACAGCATGTATCCAGGGCTTACACCATGAGGAGGACTAG	2010
Db	2505	GCAACACTTCACTTATTTCAATATACATGATTCAGAGGCTTACACGATGAGGAGGACTAG	2563
RESULT 5			
AR208616			
LOCUS	AR208616	2532 bp	DNA linear PAT 20-JUN-2002
DEFINITION	Sequence 4 from patent US 6383778.		
ACCESSION	AR208616		
VERSION	AR208616.1	GI:21509816	
KEYWORDS	Unknown.		
SOURCE	Unknown.		
ORGANISM	Unclassified.		
REFERENCE	1 (bases 1 to 2532)		
AUTHORS	Zuker C.S., Adler J.E. and Lindemeier J.		
TITLE	Nucleic acids encoding a G-protein coupled receptor involved in		
JOURNAL	sensory transduction		
FEATURES	Patent: US 6383778-A 4 07-MAY-2002;		
source	Location/Qualifiers		
	1..2532		
BASE COUNT	/organism="unknown"		
ORIGIN	498 a 824 c 631 g 579 t		
Query Match	60.3%; Score 1212.6; DB 6; Length 2532;		
Best Local Similarity	76.3%; Pred. No. 2.1e-190;		
Matches 1555; Conservative	0; Mismatches 454; Indels 30; Gaps 4;		
QY	2 TCACCTACAGGGCCATCAGCATGAGTGCAGACAGAGTGGCGCTTCCCGGCTTGTCTGC 61		

Db 494 TCACATATAGGGCCATCACGCAACAAGCTGCAAGACAGGGGGCTTCCTGCTGCACTGCG 553
 QY 62 GTACACACCCAGCGCGAGACCAACAGTGCAGGCCATGCTGACGTGATGCTGCACTTCC 121
 Db 554 GCACGTGCGCCGAGCGCCACCCACACATCAGAGGCCATGCTGACATGTATGCTTCACTTCC 613
 QY 122 GCTGAACATGATCTATGCTGCTGAGCAGAGACACTATGCGCGGCAACAAATGCGCAGC 181
 Db 614 AGTGAACATGATCTGCTGCTGCTGAGCAGATGACGATTTATGCGCGAAGAACACCGCACC 673
 QY 182 TGCCTGAGGAGCGCGTGC---CCGGCGGACATCTGATGCTGCTTCCAGAGAGAGCTGC 238
 Db 674 TGTGAGCGAGGCTGTGACCAACATCGGACATCTGATGCTTCCAGAGAGCTTGC 733
 QY 239 CCACACTGACGCCCAACCAAGAACATGAGCTGAGAGAGCGCCGCTGCTGACCATTTG 298
 Db 734 CCGTACCAAGAACCCCAACAGAGCTGTGAGGCTGTGAGAGACAGAACCACTGTGCAACATCC 793
 QY 299 TGGACAACTGACAGAGACAGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 358
 Db 794 TGGACAACTGAGGGGAGCTTGGGGCGCTGTGTGTGTATATCTTCCGCGAGCTGAGCG 853
 QY 359 TGTACCACTTCTTCAATAGGTGCTGCGCAAGAACTTCAAGCGCGCGCTGTGTGTGCT 418
 Db 854 TGCACAACTTCTTCCGTGAGGTGCTGCGCTGGAACCTTCAAGCGCGCTTGTGTGTGATTTGCT 913
 QY 419 CCGAGTCTGAGGGCATGACCGCGCTGCTGCAACACTTCAAGAGAGTGGCGCACTTGGGCA 478
 Db 914 CTGATCTGAGGCGCTGACACCTGCTTCTTCAACAACTGACAGAGTGTGCGCACACAGGGCA 973
 QY 479 CTTTCTGAGGATCACATTCACAGAGCTGTGCCATCCGGGCTTCAAGATTTCCGAGAT 538
 Db 974 CTTTCTGAGGATCACATTCACAGAGAGGTGTCCATCCCTGCTTCAAGAGTTCCGAGAT 1033
 QY 539 GGGGCGCACAGAGCTGAGCG 598
 Db 1034 GCGATGACCAACCGAGAGGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1093
 QY 599 AGGAGTGGCAACAGTGGCTGAAAGCGCACTTGTCTTCAACAACTTCAAGAGTGGCTGCTG 658
 Db 1094 AGGAGTGGCAACAGTGGCTGAAAGCGCACTTGTCTTCAACAACTTCAAGAGTGGCTGCTG 1153
 QY 659 GGGAGCGTGTGCTTACAGCGGTGTACTGTGCGGTGTATGCTGTGCGCGCGCGCGCGCGCA 718
 Db 1154 GGGAGCGTGTGCTTACAGCGGTGTACTGTGCGGTGTATGCTGTGCGCGCGCGCGCGCGCA 1213
 QY 719 GCTTCTGCTGCTGTGACAAAGACCTGACCAAGAGGAGGTGTACTGCTTGGCAGCTGC 778
 Db 1214 GACTCTCTCAGCTGCAATAGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 1273
 QY 779 TTGAGGAGATCTGGAAGGTCAACTTCACTTCTGAGACCAACAACTTCTTCCAGCCGC 838
 Db 1274 TCAGGAGATCTGAGATCTGCACTTCAAGCTTGTGCGCGCGCGCGCGCGCGCGCGCGCG 1333
 QY 839 AAGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 898
 Db 1334 AAGGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1393
 QY 899 CTTTCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 955
 Db 1394 CTTTCAAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1453
 QY 956 CATCTCTGCAACCGTCAACCAACAGATCCCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1015
 Db 1454 TGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1513
 QY 1016 CAGGGCAAG 1075
 Db 1514 CTGGGCAATGAG 1573
 QY 1076 TTCCCGGACCTTCTTCAACCAACACTGAA-----TGCCCGAATA 1114
 Db 1574 CGCGGAGACCTTCAACCAACAGATGAGATGAGATTTAAGTGTGCTGCTGCGCGGCTT 1633

QY 1115 ACGATGCTCTTACCAAGTGAAGACTCTCTGCTTCAAGCGGAGAGTGGTCTTCTGGAAT 1174
 Db 1634 CCATGTGGCTTCAAG 1693
 QY 1215 GGCATGAG 1234
 Db 1694 GGCATGAG 1753
 QY 1235 TGGCCATCTGCTGATATTTGAG 1294
 Db 1754 TGGCCATCTGCTGATCTTCTGAG 1813
 QY 1295 GCGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1354
 Db 1814 GCGCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1873
 QY 1355 ACGTGGGCGCGCGAG 1414
 Db 1874 ATGTGGGCGCGCGAG 1933
 QY 1415 TCACAAATTTGATCTCTGATATGCGCGGTGCTTCTTCCAGATGCTGCGCTTCAAGA 1474
 Db 1934 TCTCGTCTGCTCTCTGATATGCGCGGTGCTTCTTCCAGATGCTGCGCTTCAAGA 1993
 QY 1475 TGGCGAGCGGCTTCCAGAGCGGCTGACGCTACTGAGGTCCGCTACAGAGCGGCTTACGCT 1534
 Db 1994 TGGCGAGCGGCTTCCAGAGCGGCTTACGCTACTGAGGTCCGCTACAGAGCGGCTTACGCT 2053
 QY 1535 CTATGCAATTTATCAAGGTACTCAAAATATGCTATGCTATGCTATGCTATGCTATGCTATGCT 1594
 Db 2054 TGTGCGCTTATCAAGCGGCTTCAAGGTACTCAAAATATGCTATGCTATGCTATGCTATGCTATGCT 2113
 QY 1595 CTGAGTCCGAGCG---CCGCTACTGAGCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1651
 Db 2114 CCATCAACCGCGAG 2173
 QY 1652 ACCCAATCAAGCGAG 1711
 Db 2174 ACCCAATCAAGCGAG 2233
 QY 1712 TGGGTTTCAAGCTTCCGCTACATGAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1771
 Db 2234 TGGGTTTCAAGCTTCCGCTACATGAGGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2293
 QY 1772 TCATACCTTCAAGATGAGCTTCTTATTCACCTATTCGCTCTGCTGACCTTCAATG 1831
 Db 2294 TCATACCTTCAAGATGAGCTTCTTATTCACCTATTCGCTCTGCTGACCTTCAATG 2353
 QY 1832 CTGCGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1891
 Db 2354 CTGCGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2413
 QY 1892 TGGCCATGAGCTTGGCTACTTCCG 1951
 Db 2414 TGGCCATGAGCTTGGCTACTTCCG 2473
 QY 1952 GCACACGCGCGCTTCAACAGAGATGATCAGAGGCTTACACCATGAGAGAGAGAGAGAGAGAGAG 2010
 Db 2474 GCACACGCGCGCTTCAACAGAGATGATCAGAGGCTTACACCATGAGAGAGAGAGAGAGAGAGAG 2532
 RESULT 6
 AT032623
 LOCUS 2532 bp mRNA linear ROD 05-FEB-2002
 DEFINITION Mus musculus sweet taste receptor T1R2 mRNA, complete cds.
 ACCESSION AY032623
 VERSION AY032623.1 GI:15147680
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1475 TGGCAGCGGCTGCGCCAGCGGCTACAGTACGAGGCGGCTACGAGGCGGCTACGAGT 1534
 1994 TGGCCAGACGCGCTGCGCAGCGGCTACGAGTACGAGTACGAGGCGGCTACGAGT 2053
 1535 CTATGAGCTTATTCAGCGGCTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGT 1594
 2054 TCGTGGCGCTTATTCAGCGGCTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGT 2113
 1595 CTGAGTCCAGCC---CCGTAAGTACCGGCTACGAGTACGAGTACGAGTACGAGTACGAGT 1651
 2114 CCAACACCGCTATGCGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGG 2173
 1652 ACCCAATACCGGAGCGGCTGCTGCTACAGCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 1711
 2174 ACCCAATACCGGAGCGGCTGCTGCTACAGCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2233
 1712 TGGGTTGAGCTGCGCTACATGAGGCAAGAGCTGCGCAGCAACATACAGAGCGCAAGT 1771
 2234 TGGGTTGAGCTGCGCTACATGAGGCAAGAGCTGCGCAGCAACATACAGAGCGCAAGT 2293
 1772 TCATCAGCTTACAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGT 1831
 2294 TCATCAGCTTACAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGT 2353
 1832 CTGCTACAGCGGCGGCTGCTGCTACAGCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1891
 2354 CTGCTACAGCGGCGGCTGCTGCTACAGCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2413
 1892 TGGCAGTACGCTGCGCTACATGAGGCAAGAGCTGCGCAGCAACATACAGAGCGCAAGT 1951
 2414 TGGCAGTACGCTGCGCTACATGAGGCAAGAGCTGCGCAGCAACATACAGAGCGCAAGT 2473
 1952 GCAACAGCGGCGGCTGCTGCTACAGCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2510
 2474 GCAACAGCGGCGGCTGCTGCTACAGCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2532

RESULT 7
AF458149S6
LOCUS AF458149S6 930 bp DNA linear PRI 12-APR-2002
DEFINITION Homo sapiens taste-specific G-protein coupled receptor T1R2
ACCESSION AF458154
VERSION AF458154.1 GI:20142332
KEYWORDS AF458154.1 GI:20142332
SEGMENT 6 of 6
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE 1 (bases 1 to 930)
AUTHORS Li, X., Staszewski, L., Xu, H., Durick, K., Zollner, M. and Adler, E.
JOURNAL Human receptors for sweet and umami taste
PROC. Natl. Acad. Sci. U.S.A. 99 (7), 4692-4696 (2002)
PUBMED 21927605
REFERENCE 2 (bases 1 to 930)
AUTHORS Li, X., Staszewski, L. and Adler, E.
TITLE Direct Submission
JOURNAL Submitted (14-DEC-2001) Genomix, Inc., 11099 N. Torrey Pines Road,
 La Jolla, CA 92037-1007, USA
FEATURES location/Qualifiers
 1..930
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 order(AF458149.1:1..184,AF458150.1:1..305,
 AF458151.1:1..774,AF458152.1:1..210,AF458153.1:1..123,
 1..930)
 /gene="TAS1R2"
 /note="T1R2"
 join(AF458149.1:1..184,AF458150.1:1..305,
 AF458151.1:1..774,AF458152.1:1..210,AF458153.1:1..123,
 1..930)

AF458151.1:1..774,AF458152.1:1..210,AF458153.1:1..123,
 1..930)
 /gene="TAS1R2"
 /product="taste-specific G-protein coupled receptor T1R2"
 join(AF458149.1:1..184,AF458150.1:1..305,
 AF458151.1:1..774,AF458152.1:1..210,AF458153.1:1..123,
 1..930)
 /gene="TAS1R2"
 /note="T1R2: putative human sweet taste receptor"
 /codon_start=1
 /product="taste-specific G-protein coupled receptor T1R2"
 /protein_id="AA012233.1"
 /db_xref="GI:20142333"
 /translation="MGRPRKICSLFLLWLAIPAENSDPFLPGDYLGLFSLHAN
 MKGIVHLNLOVPKREYEVYICVNLNMOARFVERINDSSILPVLIGYELFVHC
 YISNNVPLVFLTHEDNLEPIQEDISYISRAVAVIGPDNSESVMVAVNLSFLPLP
 QILYSALISDELDRDVRPALLRTTPSADHVEAVQMLHFRMMLIIVLVSDYVGRD
 NGOLIGERVARRDICIAFOETPLIQNONMSEERDLVYDKLQDSTARVYVPS
 PDLTYHFNENFLRNFETGAVWIASMAIDPVLANTELGHTFLGTITQSPVIG
 FSEEREMGPOAGPPPLSRTSOSYCNOCNCLNATLSPNTILSLGSRVYVYSAY
 YAVAHALSHLGGCKSTCKRYVYPMOLBEIKVNTLIDHOIFPDPOGVALHLET
 VOMQWDRSQNFQSVASTYPLQROLKNIODISHTVNTTIPMSKSRKSGOKKFPV
 GIHYCEECIDCLPGLNTHEDYECQACPNMMSYQSETSCEROLVLEMEHAPL
 IAVALLAAGLSTLAILVIFMRHPTPIYASAGPMCFMLTLVAVAVYVYVGP
 PKVSTLCROALFPLCFCTICISIAVSFOYCAFKMSPRAYSVWRQGYVSM
 AFITVLKMLVIVIGMLATGLSPTRPDQDKITVGCNPNRSLFNTSLDILSV
 VGSFPAIMGKRLPNTYNAKRTITSMFTYFISVSLCTPMKATYGLVTLVYDLVTVL
 NLAISLGYEPKCYMILFPERNTPAFVSMIGYTRRD"
 1..930
 /gene="TAS1R2"
 /number=6
 BASE COUNT 171 a 322 c 222 g 215 t
 ORIGIN
 Query Match 43.8%; Score 879.8; DB 9; Length 930;
 Best Local Similarity 99.4%; Pred. No. 1.8e-133;
 Matches 904; Conservative 0; Mismatches 2; Indels 3; Gaps 2;

1105 TGGCCGAATACGAGTGGCTCCACAGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGT 1164
 22 TGGCCGAATACGAGTGGCTCCACAGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGT 81
 1165 TTCCGTGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGT 1224
 82 TTCCGTGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGTACGAGT 141
 1225 CTGACGACCCCTGCGCATCTGCTGATATTCGAGGACCTTCGACAGACCCCAATAGTTCGC 1284
 142 CTGACGACCCCTGCGCATCTGCTGATATTCGAGGACCTTCGACAGACCCCAATAGTTCGC 201
 1285 TTGGGTGGGGGGCCCATGCTGCTCTGATGCTGACACACGCTGCTGCTGCTGCTGCTGCTGCTGCT 1344
 202 TGGGTGGGGGGCCCATGCTGCTCTGATGCTGACACACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 261
 1345 GTCCCGGTGATGAGTGGGGGGCCCAAGGCTTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1404
 262 GTCCCGGTGATGAGTGGGGGGCCCAAGGCTTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 321
 1405 CCGCTGCTGCTACAAATTTGATCTGCTGATGCGCGTGGCTGCTTTCACAGATGCTTCGC 1464
 322 CCGCTGCTGCTACAAATTTGATCTGCTGATGCGCGTGGCTGCTTTCACAGATGCTTCGC 381
 1465 GCGTTCAGATGCGCAGCGGCTTCCAGCGGCTACGCTACGCTATGCGTCCGACAGAGG 1524
 382 GCGTTCAGATGCGCAGCGGCTTCCAGCGGCTACGCTACGCTATGCGTCCGACAGAGG 441
 1525 CCGTACGCTCTATGAGCTTATTCACAGTACGCTACGCTACGCTACGCTACGCTACGCTACGCTACG 1584
 442 CCGTACGCTCTATGAGCTTATTCACAGTACGCTACGCTACGCTACGCTACGCTACGCTACGCTACG 501
 1585 CTGGC--ACGGCTTAGTCCACG--CCGCTACGACCCCGATGACCCCAAGATCAACAATT 1641
 502 CTGGCAGCGGCGCTACGCTACGCTACGCTACGCTACGCTACGCTACGCTACGCTACGCTACGCTACG 561

QY 1642 GTCTCTGTAAACCCCACTACCGCAACAGCCTGTCTTTCAACACCACTGGACCTGCTG 1701
 DB 562 GTCTCTGTAAACCCCACTACCGCAACAGCCTGTCTTTCAACACCACTGGACCTGCTG 621
 QY 1702 CTCTCACTGTGGTTGCTTACCTTGGCTTACATGGGCAAGAGCTGCCACCACTACAAC 1761
 DB 622 CTCTCACTGTGGTTGCTTACCTTGGCTTACATGGGCAAGAGCTGCCACCACTACAAC 681
 QY 1762 GAGGCCAAGTTCAATCAACCCCTACAGATGACCTTATTTACCTACCTACCTGCTCCCTG 1821
 DB 682 GAGGCCAAGTTCAATCAACCCCTACAGATGACCTTATTTACCTACCTACCTGCTCCCTG 741
 QY 1822 ACCTTATGTCTGCTTACAGCGGGGTGCTGTCACCATGTCGACCTTGGTCACTGTG 1881
 DB 742 ACCTTATGTCTGCTTACAGCGGGGTGCTGTCACCATGTCGACCTTGGTCACTGTG 801
 QY 1882 CTCAACCTCTGGCCATCAACCTTGGCTTACATGGGCAAGAGCTGCCACCACTACAAC 1941
 DB 802 CTCAACCTCTGGCCATCAACCTTGGCTTACATGGGCAAGAGCTGCCACCACTACAAC 861
 QY 1942 TACCCGAGGCAACACCGCCCTTACTTCAACAGCATGATCCAGGCTACACCATGAG 2001
 DB 862 TACCCGAGGCAACACCGCCCTTACTTCAACAGCATGATCCAGGCTACACCATGAG 921
 QY 2002 AGGACTAG 2010
 DB 922 AGGACTAG 930

RESULT 8
 AL831755 209676 bp DNA linear HTG 09-AUG-2002
 LOCUS Homo sapiens chromosome 1 clone RP13-279N23, *** SEQUENCING IN
 DEFINITION
 ACCESSION AL831755
 VERSION AL831755.6 GI:22204771
 KEYWORDS HTG; HTGS_PHASE2; HTGS_ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 209676)
 Harrison, E.
 Direct Submission
 Submitted (02-AUG-2002) Wellcome Trust Sanger Institute, Hinxton,
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
 humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 on Aug 11, 2002 this sequence version replaced gi:21998292.
 COMMENT
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk
 ----- Project Information
 Center project name: DB279N23
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Sequencing vector: plasmid; L08752; 100% of reads
 Chemistry: Dye-terminator; 97% of reads
 Chemistry: Dye-terminator Big Dye; 2% of reads
 Consensus quality: 209374 bases at least Q40
 Consensus quality: 209478 bases at least Q30
 Insert size: 209676; sum-of-contrigs
 Insert size: 218959; 3.2% error; agarose-tp
 Quality coverage: 12.24x in Q20 bases; sum-of-contrigs Quality
 Coverage: 11.72x in Q20 bases; agarose-tp

* NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES
 source
 location/Qualifiers
 1..209676
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="1"
 /clone_11b="RP13-279N23"
 /clone_11b="RP13-279N23"
 1..209676
 /note="assembly-fragment:03689
 clone_end:SP6
 vector_side:right"
 BASE COUNT 52994 a 53379 c 52820 g 50483 t
 ORIGIN
 Query Match 43.7%; Score 878.2; DB 2; Length 209676;
 Best Local Similarity 99.3%; Pred. No. 1.7e-135;
 Matches 903; Conservative 0; Mismatches 3; Indels 3; Gaps 2;
 QY 1105 TGCCTGATATGAGAGTGGTCTTACAGAGTGAAGTCTGCTTCAAGCGGAGCTGCTG 1164
 DB 193533 TGCCTGATATGAGAGTGGTCTTACAGAGTGAAGTCTGCTTCAAGCGGAGCTGCTG 193474
 QY 1165 TTCTGGAATGGATGAGGACCAACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1224
 DB 193473 TTCTGGAATGGATGAGGACCAACCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 193414
 QY 1225 CTCAGCACCTTGGCCATCTGTGATATTTCTGAGGACATTCAGACACCATAGTTCCG 1284
 DB 193413 CTCAGCACCTTGGCCATCTGTGATATTTCTGAGGACATTCAGACACCATAGTTCCG 193354
 QY 1285 TCGGCTGGGGGGCCCATGCTTCTGATGCTGACATGCTGCTGCTGCTGCTGCTGCTGCTG 1344
 DB 193353 TCGGCTGGGGGGCCCATGCTTCTGATGCTGACATGCTGCTGCTGCTGCTGCTGCTGCTG 193294
 QY 1345 GTCCCGGTGTACGTTGGGGGGCCCAAGGCTTCCACCTGCTGCTGCTGCTGCTGCTGCTGCT 1404
 DB 193293 GTCCCGGTGTACGTTGGGGGGCCCAAGGCTTCCACCTGCTGCTGCTGCTGCTGCTGCTGCT 193234
 QY 1405 CCCCTCTGCTTCAACAAATTTGATCTCTGTATGCGCGGTCTTCCAGATCTGCTGCG 1464
 DB 193233 CCCCTCTGCTTCAACAAATTTGATCTCTGTATGCGCGGTCTTCCAGATCTGCTGCG 193174
 QY 1465 GCCTTCAAGATGGCCAGCCCTTCCACGCGCTTACAGCTACGTTGGGCTCCCTACAGGGG 1524
 DB 193173 GCCTTCAAGATGGCCAGCCCTTCCACGCGCTTACAGCTACGTTGGGCTCCCTACAGGGG 193114
 QY 1525 CCCCTACCTCTTATGAGCTTTATCAGGTAAGTCAAAAATGTCATTGTAATTTGGCATG 1584
 DB 193113 CCCCTACCTCTTATGAGCTTTATCAGGTAAGTCAAAAATGTCATTGTAATTTGGCATG 193054
 QY 1585 CTGGC--ACGGCTCAGTCTCCACC--CCGTAATGACCCCGATGACCCCAAGATCACAATT 1641
 DB 193053 CTGGC--ACGGCTCAGTCTCCACC--CCGTAATGACCCCGATGACCCCAAGATCACAATT 192994
 QY 1642 GTCTCTGTAAACCCCACTACCGCAACAGCCTGTCTTTCAACACCACTGGACCTGCTG 1701
 DB 192993 GTCTCTGTAAACCCCACTACCGCAACAGCCTGTCTTTCAACACCACTGGACCTGCTG 192934
 QY 1702 CTCTCACTGTGGTTGCTTACCTTGGCTTACATGGGCAAGAGCTGCCACCACTACAAC 1761
 DB 192933 CTCTCACTGTGGTTGCTTACCTTGGCTTACATGGGCAAGAGCTGCCACCACTACAAC 192874
 QY 1762 GAGGCCAAGTTCAATCAACCCCTACAGATGACCTTATTTACCTACCTACCTGCTCCCTG 1821
 DB 192873 GAGGCCAAGTTCAATCAACCCCTACAGATGACCTTATTTACCTACCTACCTGCTCCCTG 192814
 QY 1822 ACCTTATGTCTGCTTACAGCGGGGTGCTGTCACCATGTCGACCTTGGTCACTGTG 1881
 DB 192813 ACCTTATGTCTGCTTACAGCGGGGTGCTGTCACCATGTCGACCTTGGTCACTGTG 192754
 QY 1882 CTCAACCTCTGGCCATCAACCTTGGCTTACATGGGCAAGAGCTGCCACCACTACAAC 1941
 DB 192753 CTCAACCTCTGGCCATCAACCTTGGCTTACATGGGCAAGAGCTGCCACCACTACAAC 192694

OY	1942	TACCCGGAGCGCAACGCCGCGCTACTTCAACACATGATTCAGGGGTAAOCATTGAAG	2001
Db	192633	TACCAGGAGCCCAACACGCCCGCTACTTCAACACATGATTCAGGGGTAAOCATTGAAG	192634
OY	2002	AGGCAGTAG	2010
Db	192633	AGGCAGTAG	192625
RESULT	9		
AF458149S3			
LOCUS		774 bp	DNA
DEFINITION		(TASIR2) gene, exon 3.	linear PRI 12-Apr-2002
ACCESSION			
VERSION		AF458151	
KEYWORDS		AF458151.1	GI:20142329
SEGMENT			
SOURCE			
ORGANISM		Homo sapiens.	3 of 6
		Homo sapiens	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE			
AUTHORS		Li,X., Staszewski,L., Xu,H., Durick,K., Zoller,M. and Adler,E.	
JOURNAL		Human receptors for sweet and umami taste	
MEDLINE		Proc. Natl. Acad. Sci. U.S.A. 99 (7), 4692-4696 (2002).	
PUBMED		21927605	
REFERENCE		11917125	
AUTHORS		Li,X., Staszewski,L. and Adler,E.	
TITLE		Direct Submission	
JOURNAL		Submitted (14-DEC-2001) Genomix, Inc., 11099 N. Torrey Pines Road,	
		La Jolla, CA 92037-1007, USA	
FEATURES			
source		Location/Qualifiers	
		1..774	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/chromosome="1"	
exon		1..774	
		/gene="TASIR2"	
		/number=3	
BASE COUNT		145 a	268 c
ORIGIN			222 g
			139 t
Query Match		38.5%; Score 774;	DB 9; Length 774;
Best Local Similarity		100.0%;	Pred. No. 5.1e-118;
Matches 774;		Conservative 0;	Mismatches 0;
			Indels 0; Gaps 0;
OY	1	ATACCTCAGCGCATAGGATGATGGTGAGACAAGTGCCTTCCCGCTTTGGTG	60
Db	1	ATCACCTAAGGGCCATAGCGATGATGGTGAGACAAGTGCCTTCCCGCTTTGGTG	60
OY	61	GCTACACACCCAGCGCCGACACCAACCACTCGAGGCGATAGTGACATGCTGACATTC	120
Db	61	GCTACACACCCAGCGCCGACACCAACCACTCGAGGCGATAGTGACATGCTGACATTC	120
OY	121	CGCTGGAACGTGATTTGTGTGTGTGAGCAGCGACACTTATGCGCGGACAAATGGCCAG	180
Db	121	CGCTGGAACGTGATTTGTGTGTGTGAGCAGCGACACTTATGCGCGCGGACAAATGGCCAG	180
OY	181	CTGCTTGGCAGACGGTGGCCGCGGCGACAACTTCATGCGCTTCAGAGAGACGTGGCC	240
Db	181	CTGCTTGGCAGACGGTGGCCGCGGCGACAACTTCATGCGCTTCAGAGAGACGTGGCC	240
OY	241	ACACTGACCCCAACAGAACATGACGTGAGAGAGCGCCACGCGCTGTGACATTGTG	300
Db	241	ACACTGACCCCAACAGAACATGACGTGAGAGAGCGCCACGCGCTGTGACATTGTG	300
OY	301	GACAACTGACGAGAGAGACAGCGCGGCGTGTGCTGCTGCGCCGACCTGACCTG	360
Db	301	GACAACTGACGAGAGAGACAGCGCGGCGTGTGCTGCTGCGCCGACCTGACCTG	360
OY	360	GACAACTGACGAGAGAGACAGCGCGGCGTGTGCTGCTGCGCCGACCTGACCTG	360
Db	360	GACAACTGACGAGAGAGACAGCGCGGCGTGTGCTGCTGCGCCGACCTGACCTG	360

Oy	361	TACCACTTCTTCATAGAGGTGCTGGGCCAAGAATTACAGGGGGCCGTGGATCGCTTC	420	
Dd	361	TACCACTTCTTCATAGAGGTGCTGGGCCAAGAATTACAGGGGGCCGTGGATCGCTTC	420	
Oy	421	GAGTCCTGGGCATGACCCTGGTCTCTGCACAACAACCTCAGAGACTTGGGCACCTTGGGCACC	480	
Dd	421	GAGTCCTGGGCATGACCCTGGTCTCTGCACAACAACCTCAGAGACTTGGGCACCTTGGGCACC	480	
Oy	481	TTCTCTGGGCATTCACATCCAGAGCGTGGCCATCCGGGCTTCAGTAAGTTCCGGCAGATGG	540	
Dd	481	TTCTCTGGGCATTCACATCCAGAGCGTGGCCATCCGGGCTTCAGTAAGTTCCGGCAGATGG	540	
Oy	541	GGCCCCACAGGGCTGGGGCGCGCACCCCTTCAGAGGAGCACACGACAGGTATAACTGCAAAACAG	600	
Dd	541	GGCCCCACAGGGCTGGGGCGCGCACCCCTTCAGAGGAGCACACGACAGGTATAACTGCAAAACAG	600	
Oy	601	GAGTCGCAACATGCTGCTGACAGCGCACCTTGTCTCTTCAACACCATCTCAGGCTCTCTGGG	660	
Dd	601	GAGTCGCAACATGCTGCTGACAGCGCACCTTGTCTCTTCAACACCATCTCAGGCTCTCTGGG	660	
Oy	661	GAGCGTGTCTGTACAGCGTGTACTCTGCGGTCTATGCTGTGGGCCATAGCCTTGACACAGC	720	
Dd	661	GAGCGTGTCTGTACAGCGTGTACTCTGCGGTCTATGCTGTGGGCCATAGCCTTGACACAGC	720	
Oy	721	CTCCCTGGGCTGTGACAAAAGCACTGACACCAAAGAGGTGTCTACCCCTGGCAG	774	
Dd	721	CTCCCTGGGCTGTGACAAAAGCACTGACACCAAAGAGGTGTCTACCCCTGGCAG	774	
RESULT 10				
ALB31790		206074 bp DNA linear HTG 17-AUG-2002		
LOCUS		Mus musculus chromosome 4 clone RP23-334I5, *** SEQUENCING IN		
DEFINITION		PROGRESS ***, in ordered pieces.		
ACCESSION		ALB31790.2 GI:22416181		
VERSION		HTG; HTGS_PHASE2; HTGS_ACTIVERFIN; HTGS_DRAFT; HTGS_FULLTOP.		
KEYWORDS		house mouse.		
SOURCE		Mus musculus		
ORGANISM		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE		Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
AUTHORS		1 (bases 1 to 206074)		
JOURNAL		Wall,M.		
COMMENT		Direct Submission Submitted (16-AUG-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk On Aug 21, 2002 this sequence version replaced gi:21272737. ----- Genome Center Center: Wellcome Trust Sanger Institute Center code: SC Web site: http://www.sanger.ac.uk Contact: humquerry@sanger.ac.uk ----- Project Information Center project name: bm334i5 ----- Summary Statistics Assembly program: XGAP4; version 4.5 Chemistry: dye-terminator; 100% of reads Consensus quality: 205815 bases at least Q40 Consensus quality: 206010 bases at least Q30 Consensus quality: 206049 bases at least Q20 Insert size: 206074; sum-of-contigs Insert size: 220515; 2.7% error; agarose-fp Quality coverage: 7.94x in Q20 bases, sum-of-contigs Quality coverage: 7.73x in Q20 bases; agarose-fp ----- * NOTE: This is a 'working draft' sequence. * This sequence will be replaced * by the finished sequence as soon as it is available and * the accession number will be preserved. Location/Qualifiers 1..206074		
FEATURES				
SOURCE				


```

/db_xref="taxon:10090"
/chromosome="4"
/clone="RP23-33415"
/clone_1ib="RP23-23"
misc_feature 1..206074
             /note="assembly_fragment:00142"
BASE COUNT 48675 a 52230 c 53898 g 51211 t
ORIGIN
Query Match      30.1%; Score 605; DB 2; Length 206074;
Best local Similarity 79.9%; Pred. No 2e-90;
Matches 726; Conservative 0; Mismatches 180; Indels 3; Gaps 1;

QY 1105 TGGCCGAATACGAGTGTCTTACAGAGTGAAGACCTCTCTTCAAGCGGACCTGTC 1164
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52178 TGGCCGGGATTCATGTGTCTTACAGAGTGAAGACCTCTCTTCAAGCGGACCTGTC 52237
QY 1165 TTCTGTGAATGGCATAGGACCCACCATGCTGTGCGCCCTGCGCGCCCTGCGCTTC 1224
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52238 TTCTGTGAATGGCATAGGACCCACCATGCTGTGCGCCCTGCGCGCCCTGCGCTTC 52297
QY 1225 CTCAGCACCTGGGACATCTGTGTATCTGTGAGGACATTCAGACCATATGTTGCG 1284
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52298 ATCAGTACGCTGGGACATCTGTGTATCTGTGAGGACATTCAGACCATATGTTGCG 52357
QY 1285 TCGGCTGGGGGCCCATGTCTTCTGTATCTGTGACACTGCTGTGTGAGCATATGATG 1344
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52358 TCGGCTGGGGGCCCATGTCTTCTGTATCTGTGACACTGCTGTGTGAGCATATGATG 52417
QY 1345 GTCCCGGTGTACGTGGGGGCCCATAGGTCTTCAACCTGCTTCCCGCCAGCCCTTTT 1404
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52418 GTCCCGGTGTACGTGGGGGCCCATAGGTCTTCAACCTGCTTCCCGCCAGCCCTTTT 52477
QY 1405 CCCCTGTCTTCAATTTGATCTCTGTATGCGCGTGGTCTTCCAGATGCTGTCG 1464
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52478 ACCGTTGCTTCTTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 52537
QY 1465 GCTTCAAGTGGGACCGGCTTCCACGCGGCTTACAGTGTGCTGCTGCTGCTGCTG 1524
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52538 GCTTCAAGTGGGACCGGCTTCCACGCGGCTTACAGTGTGCTGCTGCTGCTGCTG 52597
QY 1525 CCCCTGTCTTCAATTTGATCTCTGTATGCGCGTGGTCTTCCAGATGCTGTCG 1584
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52598 CCCCTGTCTTCAATTTGATCTCTGTATGCGCGTGGTCTTCCAGATGCTGTCG 52657
QY 1585 CTGGACAGGCGC---TCAGTCCACACCCGCTACTGACCCCGATGACCCCAAGATCACA 1641
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52658 CTGGACAGGCGC---TCAGTCCACACCCGCTACTGACCCCGATGACCCCAAGATCACA 52717
QY 1642 GTCTCTGTAAACCCCACTACCGGACAGCGCTGTGTCAACACACGCTGAGCTGCTG 1701
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52718 GTCTCTGTAAACCCCACTACCGGACAGCGCTGTGTCAACACACGCTGAGCTGCTG 52777
QY 1702 CTCTCTGTGTGGGTTTACCTTCTGCTTACATGAGGCAAAAGCTGCGCCACCACTAACA 1761
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52778 CTCTCTGTGTGGGTTTACCTTCTGCTTACATGAGGCAAAAGCTGCGCCACCACTAACA 52837
QY 1762 GAGGCAAGTTCATACACCTTACAGCATGACCTTATTTACCTTACCTGCTGCTGCG 1821
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52838 GAGGCAAGTTCATACACCTTACAGCATGACCTTATTTACCTTACCTGCTGCTGCG 52897
QY 1822 ACCTTCTGTGTGGGTTTACCTTCTGCTTACATGAGGCAAAAGCTGCGCCACCACTAACA 1881
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52898 ACCTTCTGTGTGGGTTTACCTTCTGCTTACATGAGGCAAAAGCTGCGCCACCACTAACA 52957
QY 1882 CTCAACCTCTGTGGGTTTACCTTCTGCTTACATGAGGCAAAAGCTGCGCCACCACTAACA 1941
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 52958 CTCAACCTCTGTGGGTTTACCTTCTGCTTACATGAGGCAAAAGCTGCGCCACCACTAACA 53017
QY 1942 TACCCGGAGGCAACAGCGCGCTTACTTCAAGCATGATGATGATGATGATGATGATGATG 2001
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 53018 TACCCGGAGGCAACAGCGCGCTTACTTCAAGCATGATGATGATGATGATGATGATGATG 53077

```

QY 2002 AGGAGCTAG 2010
 Db 53078 AAGAGCTAG 53086

RESULT 11
 AC115397/c
 LOCUS
 DEFINITION
 AC115397
 AC115397.2 GI:21737066
 HTG: HTGS_PHASE1.
 KEYWORDS
 SOURCE
 ORGANISM
 Rattus norvegicus
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.

REFERENCE AUTHORS

1 (bases 1 to 154891)
 Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Alshrocks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T.,
 Barbieri,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D.,
 Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
 Bunay,C., Burch,P., Burkett,C., Butrell,K.L., Byrd,N.C.,
 Carion,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Haylak,P., Hewes,A., Hernandez,J.,
 Hernandez,O., Hodgson,A., Hogue,M., Holloway,C., Hollins,B.,
 Homs,F., Howard,S., Huber,J., Huliy,S., Hume,J., Jackson,L.E.,
 Jacobson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Karlsson,B., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C.,
 Kratoch,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
 Li,J., Li,Z., Lichtag,O., Liu,C., Liu,J., Liu,W., Louisedge,H.,
 Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
 Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
 Massey,E., Mawhney,E., McLeod,M.P., Meadow,M., Mel,G., Metzger,M.,
 Miner,G., Miner,Z., Mitchell,T., Mohabat,K., Morgan,M., Morris,S.,
 Moser,N., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
 Nguyen,N., Nickerson,E., Nwokoko,M., Ogun,M., Okunoye,G.,
 Oragunye,N., Oyiedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
 Peters,L., Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y.,
 Rives,M., Rojas,A., Rojudoan,I., Rolfe,M., Ruiz,S., Savery,G.,
 Scherer,S., Scott,G., Shen,H., Shooshbari,N., Sisson,I.,
 Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
 Sutton,A., Svatek,A., Tabori,P., Tamerisa,A., Tamerisa,K., Tang,H.,
 Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
 Usmani,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,
 Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
 Williams,G., Williamson,A., Wleczky,R., Wooden,S., Worley,K.,
 Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
 Weinstein,G. and Gibbs,R.

TITLE JOURNAL AUTHORS TITLE JOURNAL

Unpublished
 2 (bases 1 to 154891)
 Worley,K.C.
 Direct Submission
 Submitted (19-MAR-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 3 (bases 1 to 154891)

REFERENCE AUTHORS TITLE JOURNAL

Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

QY 121 CGGTGAGTGTATGCTATGCTGTGAGACGACACCTATGCGCGCAATGCGCAG 180
 19808 CAATGAGACTGATGATGGTGTGTGAGAGACGATATAGCGCGCAACAGCAGCAG 19749
 QY 181 CTGCTTGGGAGGCGCTGGC---CGCGGACATCTGCATGCGCTTCCAGAGACGCTG 237
 19748 CTGTTTGGGAGGCGCTGGC---CGCGGACATCTGCATGCGCTTCCAGAGAGCTTCTG 19689
 QY 238 CCCACACTGACGCGCAACGACGATGAGCTGAGAGAGGCGGCGGCGGACCATT 297
 19688 CCCATACCTGAGTCCAGCCAGGTCATGAGTCCGAGAGAGAGACACCTGAGACATC 19629
 QY 298 GTGAGAGAGCTGACAGACAGACAGCGGCTGCTGCTGCTTCCGCGGACCTGAC 357
 19628 CTGAGACAGCTGCGGCGGAGCTGCGGCGGCTGCTGCTGCTTCCGCGGAGCTGAGC 19569
 QY 358 CTGTACCATCTTCTCAATGAGTGTGCTGCGCAGACACTTCAAGCGCGCTGTGATCGC 417
 19568 CTGTATAGCTTCTTCTCAATGAGTGTGCTGCGCAGACACTTCAAGCGGCTTGTGATCGC 19509
 QY 418 TCCGAGCTGCGGCGCAGACAGACAGCGGCTGCTGCTGCTTCCGCGGACCTTGGCG 477
 19508 TCTGAGTCTGCGGCTGCTGCGGCGCAGACAGCTGCTGCTGCTTCCGCGGAGCTGAG 19449
 QY 478 ACCTTCTGCGGCGCAGACAGACAGCGGCTGCTGCTGCTTCCGCGGAGCTTCCGCGAG 537
 19448 ACTTCTGCGGCGCAGACAGACAGCGGCTGCTGCTGCTTCCGCGGAGCTTCCGCGAG 19389
 QY 538 TGGGCGGCGCAGACAGACAGCGGCTGCTGCTGCTTCCGCGGAGCTTCCGCGAG 597
 19388 CGCGGCGCAGACAGACAGCGGCTGCTGCTGCTTCCGCGGAGCTTCCGCGAGCTTCCG 19329
 QY 598 CAGAGTGTGCGACAGCTGCTGAGACAGCGGCTGCTGCTTCCGCGGAGCTTCCGCGAG 657
 19328 CAGAGTGTGCGACAGCTGCTGAGACAGCGGCTGCTGCTTCCGCGGAGCTTCCGCGAG 19269
 QY 658 GGGGAGCGCTGCTGAGACAGCGGCTGCTGCTTCCGCGGAGCTTCCGCGAGCTTCCG 717
 19268 GGGGAGCGCTGCTGAGACAGCGGCTGCTGCTTCCGCGGAGCTTCCGCGAGCTTCCG 19209
 QY 718 AGCTCTGCTGCGGCTGAGACAGCGGCTGCTGCTTCCGCGGAGCTTCCGCGAGCTTCCG 777
 19208 AGCTCTGCTGCGGCTGAGACAGCGGCTGCTGCTTCCGCGGAGCTTCCGCGAGCTTCCG 19149
 QY 778 CTGAGAGAGCTGAGAGCTGAGAGCTTCCG 806
 19148 GGTCCGAACTGCGAGGCGCTGAGCTCCAC 19120
 RESULT 12
 AFI27389 2520 bp mRNA linear ROD 04-MAR-1999
 LOCUS AF127389
 DEFINITION Rattus norvegicus putative taste receptor TR1 mRNA, partial cds.
 ACCESSION AF127389
 VERSION AF127389.1 GI:4337085
 KEYWORDS
 SOURCE Rattus norvegicus.
 ORGANISM Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
 Rattus.
 1 (bases 1 to 2520)
 Hoon, M.A., Adler, E., Lindemeier, J., Battey, J.F., Ryba, N.J. and
 Zuker, C.S.
 TITLE Putative mammalian taste receptors: a class of taste-specific GPCRs
 with distinct topographic selectivity
 JOURNAL Cell 96 (4), 541-551 (1999)
 MEDLINE 99159821
 PUBMED 10052456
 REFERENCE 2 (bases 1 to 2520)
 Hoon, M.A., Adler, E., Lindemeier, J., Battey, J.F., Ryba, N.J.P. and
 Zuker, C.S.

TITLE Direct Submission
 JOURNAL Submitted (10-FEB-1999) Taste and Smell Unit, MIDCR, 10 Center
 Drive MSC 1188, Bethesda, MD 20892-1188, USA
 FEATURES Location/Qualifiers
 source 1..2520
 /organism="Rattus norvegicus"
 /strain="Wistar"
 /db_xref="taxon:10116"
 /cell_type="Circumvallate papilla"
 1..>2520
 /note="G protein-coupled receptor"
 /codon_start=1
 /product="putative taste receptor TR1"
 /protein_id="AA018069.1"
 /db_xref="GI:4337086"
 /translation="MLFMAHLLLSLQLYCMAFSCQRTSPGSLPDGFLAGLS
 LHGDLQVHRPRLVMSQDPDSFNGHGLRQAMPFVPEINSSALLPNTLGEIY
 DYCSEANVATLRLVALDQPHRIEORDLRHSSKYAFIFPDNTDAVTTALLGP
 FLMPVLYSEASSVLSAKRPFSLKTPSDRHQVAVQQLQSGWVLSLIGSD
 YQLGVQALEELAVPRGICVAFKDIIVFASRGDPRMSMOHLQARTVVVFSNR
 HLARFERSVLANLTKVWVAEEMALSTYITSTYIGIOTLVGAQVQVPEGLK
 EEESEYRAVTAAPSACPEGSMCTNOLRECHETPTLRMTLGAESMAAYRVEAV
 VAVAGHLLIGCTSEICRGPRVPMQLQIYKYNLHENTVAFPDNDGLGYDI
 IAMDNGEPMTEIIGSASLSVHLDINKTQIOMGNKNQVYVSTCTDCLGHRVY
 VSHHCPEVCEPCEGTEFLMSLHLCQCGTEVNAKNSSTICFPRVEFLMHEPIS
 IVLIANLTLILLVGTAGLFAMHHTPVVRAGRLCLMLGSLVAGSCSYSEGE
 PVPACLLRQPLFSLGFALESLTIRSPOLYIIRFSTKVPFTFRTMONGAGLTV
 IYSSVTHLICTLWLVMPRTREVRQVRLNPSVIAETCEVNSGFLAETHNLLIS
 TPVCSICKELPENNEAKVCFESLLINPVSIAPFTMASIYOGSYLPAVNVLAGLTT
 LSGFSGYLPLPCVYLICRPELNNEHPOASIODYTRRCGTT"
 BASE COUNT 479 a 749 c 676 g 616 t
 ORIGIN
 Query Match 19.5%; Score 392.8; DB 10; Length 2520;
 Best local Similarity 51.4%; Pred. No. 3.2e-55;
 Matches 1046; Conservative 0; Mismatches 942; Indels 48; Gaps 4;
 QY 2 TCACCTACAGCGCATGAGCGATGAGCTGCGAGACAGAGTGGCGCTTCCGCTTGTCTG 61
 497 TCAGCTATGAGCGACAGCGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 556
 QY 62 GTACACACACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 121
 557 GTACCGTCCCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 616
 Db 122 GCTGAGACTGATCATGCTGTGCTGTGAGACAGGACACCTATGAGCGCGGCGGCGGCGGCGG 181
 617 GGTGGGTGTGATCTGCTGCTATGTGCGATGAGTATTAAGGCGGAGCTGGGTGCGAG 676
 QY 182 TGCTTGGCGAGCGGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 241
 677 CGCTGAGAGAGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 734
 QY 242 CACTGACGCCACGACGACGATGACGACGAGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 301
 735 -----TTTCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 778
 QY 302 ACAAGCTGACGACGACGACGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361
 779 AGCATCTGGCTGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 838
 QY 362 ACCACTTCTTCAATGAGTGTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 421
 839 GAGTGTCTTCAATGAGTGTGCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 898
 QY 422 AGTCCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 481
 899 AAGACTGGGCGCATCTCCAGCTACATACACGCGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 958
 QY 482 TCTGCGGCGATCAGTACAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 541
 959 TGCTCGGTGTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1018

QY 542 GCCCAGAGGCTGGGCGGCCACCCCTCAGACAGACGACGAGCTATACCTGC-----A 595
 Db 1019 ATGTCAGAGGCTGTACAGACTCTCTCCACGCTTGCCCGGAGGGTCCCTGGTGCAGCCTA 1078
 QY 596 ACCAGAGGTGGACAACTGCTTGAACGCCACTTGTCTTCAACACATCTCTCAGGCTCT 655
 Db 1079 ACCAGAGGTGGCGGAGAGGCCACAGCTTACAGACTCTGTACATGCCCCACGGCTTGAAGCT 1138
 QY 656 CTGGGAGAGCTGTGCTACAGAGCTGTACTCTGGGGTCTATGCTGTGGCCCATGCCCTGC 715
 Db 1139 TCTCATAGATGCGCGCTTACAGAGTGAATGAGCTGTGTAGCTGTGGCCACGCGCTCC 1198
 QY 716 ACAGCTCTCTGCTGGTGTACAAAGACACCTGTACACCAAGAGGGTGTACCCCTGGAGC 775
 Db 1199 ACCAGCTCTCTGGGATGTACTCTGTAGATCTGTTCAGAGGCCCATGTACCCCTTGGCAGC 1238
 QY 776 TGTGTGAGAGATCTGGAAGGTCAACTTCACTCTCTGGAGCACCAATCTTCTTGCACC 835
 Db 1259 TTTCTACAGATCTACAAAGGTGAATTTTCTTACATGAGAATACGTGGCATTTGATG 1318
 QY 836 CGCAAGGGAGCTGGCTCTGTGACTTGGAGATTTGCCAGTGGCATGGAGCCGAGCCAGA 895
 Db 1319 ACAAGGGGAGACTGTAGTACTTACGACATCATCTGCTGGAGCTGGAATGGACCTGAT 1378
 QY 896 ATCCCTTCCAGAGCGTGGCTCTACTACTACCCCTGACGACGACGTGAAGAATCAAGA 955
 Db 1379 GGACTTGTGAGATCATTTGGCTGTGCTCTACCTGTCCAGTTCACTGTGACATTAATAGA 1438
 QY 956 CATCTCT-----GCACACGCTCAACAAACAGATCCCTATGTCCATGTGTTCCAAAGGT 1009
 Db 1439 CAAAATTCAGATGCGACGGGAGAAACATCAGTGGCTGTGAGTGTGACACGAGACT 1498
 QY 1010 GCCAGTAGGCGAAGAAAGAGAGCGCTGGGAGATCCAGCTGTGCTGCTCCAGAGCAACG 1069
 Db 1499 GTCTGGCAGGGCACACAGGGTGTGTGGGTTCGCCACACTGTGCTTGTAGTGTGTC 1558
 QY 1070 ACTGCTTCCCGGACCTTCTCTCAACCACTGAA-----TGCCCGA 1111
 Db 1559 CCTGCAAGAGCTGGGAGCTTCTCAACATGATGAGCTTCACATCTGCCAGGCTTGTGAA 1618
 QY 1112 ATTAAGATGTCTTACCAGATGAGACTTCTGCTTCAAGCGGAGCTGTCTTCTGG 1171
 Db 1619 CAGAGATATGGGACCCCAAGAGAGACACTACTGCTTCCACGACGCTGAGTCTTGG 1678
 QY 1172 AATGGATAGAGGACCAACCAATGCTGTGGCCCTGCTGGCCCTGGGCTTCTCAGCA 1231
 Db 1679 CTGGGATGAACCCATCTTGTGTGCTAATAGCAGCTAACACGCTAATTTGTGTGCTGC 1738
 QY 1232 CCTGAGCATCTGTGATATTTCTGAGGCACTTCCAGACACCAATAGTTCGTGAGCTG 1291
 Db 1739 TGTGTGGAGACGTGGGCTGTGTGGCTGTGGCATTTTCCACACACCTGTAGTCAAGCTG 1798
 QY 1292 GGGGCCCCANGTCTTCTCTGATGCTGACACTGTGCTGTGGCATACATGTTGTGCTCCGG 1351
 Db 1799 GGGGTAGGCTGTGCTCTCTCACTGTGGGTCTCTGTGGTGGGAAAGTTGACGCTTCTATA 1858
 QY 1352 TGTAGGAGGGGCGCGCAAGGTTCACACCTGCTGTGGCCGAGGCCCTTCTCCCTCT 1411
 Db 1859 GCTTCTTGGGAGGACCAAGGTGCGCGCTGTGGCTGTGGCTGTGGCTGTGGCTTCTCTCG 1918
 QY 1412 GCTTACAAATTTGATCTCTGATTCGCGCTGTCTTCTTCCAGATCTGTGCGGCTTCA 1471
 Db 1919 GGTTCGATCTTCTCTCTGCTGCTGACATCCGCTCTTCCAACTGTGATCATCTTCA 1978
 QY 1472 AGATGGCAGCGCTTCCACGCGCTACAGCTACGTGGCTCGCTACACAGGGGCGCTACG 1531
 Db 1979 AGTTTTCACCAAGGTGCCCAATTTCTACCTGAGGCGCAAAACCATGGTGCAGAGCTC 2038
 QY 1532 TCTGTAGGCAATTTATCAGGATCTCAAAATGATGATGTGTATATGGCATGCTGGGAC 1591
 Db 2039 TATTCGTATTTGATGCTCAGCTCACGGTCCATTTGTGCTCATGTCTCAGATGCTTGTATGT 2098
 QY 1592 GGCTCATGTCCACCCCGTACTGACCCCGAGTGAACCCCAAGATCACAATTTGTCTCTGTA 1651

Db 2099 GGACCCCAAGACCCACAGGGAATACAGGCTTCCCCATCTGGTGTATTTCTCAGTCA 2158
 QY 1652 ACCCAACTACCGCAACAGCTGCTGTTCACACACAGGCTGGAGCTGTCTCAATGG 1711
 Db 2159 CAGAGTCAACTCTGTAGGCTTCTGCTTGTGGCTTACCAACAAATCTCTCTCAGCA 2218
 QY 1712 TGGGTTTACGTTGGCCCTACATGGGCAAGAGCTGGCCACCAATCAACAGAGCAAGT 1771
 Db 2219 GTACCTTGTGTGACGCTACCTGTGGTAAGAACTGCCAGAGAACTATATATGAAGCAAT 2278
 QY 1772 TCATCACCTCAGATGATGATCTTATTTACCTATCCGCTTCCCTGTGACCTTCAGT 1831
 Db 2279 GTGTACCTTACGCTGTCTCTCAACTTCTGTATCTGTGATGCTCTTTCACATGGGCA 2338
 QY 1832 CTGCTACAGCGGGGTGTGTACATGTCAGACCTTGTGTGATCACTGTCTCAACCTCC 1891
 Db 2339 GCATTTACCAAGGCAAGTACCTGCTGGGTCATATGTCTGGAGGCTGACACACTGA 2398
 QY 1892 TGGCCATAGCTGGGCTACTTGGCCCAAGTGTATATGATCTTCTTACCCGGAGC 1951
 Db 2399 GCGGCGCTTACAGGCTTACTTCTCCCAAGTGTATGTATCTGTGCGCTCCAGAAC 2458
 QY 1952 GCAACACGCCCGCTACTTCAACAGATATTCAGAGGCTTACACATGAGAGGAGC 2007
 Db 2459 TCACCAATACAGAACACTTTCAGGCTCTCATTCAGAGACTACAGAGGCGCTGGCGC 2514

 RESULT 13
 AX282894 2292 bp DNA Linear PAT 02-NOV-2001
 LOCUS AX282894
 DEFINITION Sequence 3 from Patent WO0172842.
 ACCESSION AY282894
 VERSION AX282894.1 GI:1609866
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE 1
 Walke,D.W., Scoville,J., Donoho,G., Turner,C.A., Friedrich,G.C.,
 Abulin,A., Zambrowicz,B., and Sands,A.T.
 TITLE Novel human 7tm proteins and polynucleotides encoding the same
 JOURNML Patent: WO 0172842-A 3 04-0CT-2001;
 Lexicon Genetics Incorporated (US)
 FEATURES
 Location/Qualifiers
 1..2292
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 437 a 690 c 638 g 524 t 3 others
 ORIGIN
 Query Match 19.2%; Score 386.4; DB 6; Length 2292;
 Best Local Similarity 51.1%; Pred. No. 3.7e-54;
 Matches 1042; Conservative 1; Mismatches 947; Indels 48; Gaps 4;

 QY 1 ATACCTTACAGCGCATAGCAGATGAGCTGGAGACAAAGTGTCCGGCTTGTGCTG 60
 Db 265 ATTACTTATGGGCGCAGAGAGAGACGCTCAGCGGAGCGGCACTATACCTTCTCTG 324
 QY 61 CGTACCAACCCAGGCGCGGACACAGCTGAGGCGATGAGTGTGAGCTGATGTCACATTG 120
 Db 325 CGACATTCCTCCATATGACATGACAGTGTGAGACATGATGTGTCTCTGCAAGATTC 384
 QY 121 CGTGAACCTGGATCATTTGTCTGTGACAGCGACACACTATATGGCGGACAAATGGCCAG 180
 Db 385 GGGTGGACCTGATCTCTGTGTTGGACAGAGTACGACTATATGGGACACTAGGGGTGAG 444
 QY 181 CTGCTGGCAGAGCGGTGGCGCGGCGGACATCTGCAATCGCTTTCAGAGAGAGTGGCC 240
 Db 445 GCATGTGGAAGACAGGCGCAGTGTAGGGGATCTGCAATGCTTTCAGAGATCATGCTCC 504
 QY 241 ACATGTAGCCCAACAGACATGATGATGAGAGAGCGCCAGCGCTGTGACATGTG 300

Db 505 -----TTCTGCGCAGGTGGGCGATGAGAGATGAGTGCCTCATG 546
 QY 301 GACAACTGAGAGACAGACAGGGCGGTGCTGTGTTCCGCCGACCTGACCTG 360
 Db 547 CGCAGCTGAGCGAGCGGGGCGACGCTGTGTGTTTTCAGCGCGGAGTTGGCC 606
 QY 361 TACCACTTCTCAATGAGTGTCTGCGCAGAACCTTCAGGGGCGGTGTGATGCGCTCC 420
 Db 607 AGGGTGTGTTTTCAGTGCCTGTGTGACCAACCTGACTGCGAAGGTGTGGCTCA 666
 QY 421 GACTCTGGGCGATGACAGCCGGTCTGACACACCTACAGAGCTGGGCGACTGGGCGAC 480
 Db 667 GAAGCGTGGGCGCTTCACAGGACATCATCTGGGTGGCGGATTCAGGCGCATTTGGGATG 726
 QY 481 TTCTGGGCGATCAGATCCAGAGAGCTGGCCATCCGGGCTTCAGTGTGCTCCGAGTGG 540
 Db 727 GTCTGTGGGCGTGGCGATCCAGAGAGGCGTGTCTGCGCTGAAGGCGTTTGAAGAGCC 786
 QY 541 GGCCACAGGCTGGGCGCGCCACCTCAGCAGAGACAGCAGAGCTAT-----ACCTGC 594
 Db 787 TATGCCCGGCGACAGAGAGAGGCGCCCTAGGCGTTGCGACAGAGGCTCTGTGCGAGAGC 846
 QY 595 AACAGAGATGCGACAACTGCTGAAGCGCAGCTTGTCTTCAACACATTCAGAGCTC 654
 Db 847 AATCAGCTCTGACAGAGATGCGCAAGCTTTCATRCACACAGATGCGCAAGCTCAAGGCC 906
 QY 655 TCTGGGAGGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 714
 Db 907 TTTCTCATGATGTTCTGCTTCAACAGCATACCGGCTGTGTATGCGGTGCGCATGCGCTC 966
 QY 715 CACAGCTCTCTGCGCTGTGACAAAGCAGCTGACCAAGAGGCTGTCTACCCCTGGCAG 774
 Db 967 CACACACTCTCTGCGCTGTGCTGTGAGCTTGTTCAGAGGCGCGAGCTACCCCTGGCAG 1026
 QY 775 CTGCTGTGAGAGATCTGAGAGTCACTTCACTCTCTGAGACCAACATCTTCTTGAC 834
 Db 1027 CTTTGTGAGAGATCAACAAAGGTGATTTCTTCTTCAACAAAGACACTGTGGCTTAAAT 1086
 QY 835 CCGCAAGGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 894
 Db 1087 GACACAGAGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1146
 QY 895 AATCCTTCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 948
 Db 1147 TGGACCTTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1206
 QY 949 ATCAAGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1008
 Db 1207 ACCAAATTCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1266
 QY 1009 TGGCAGTCAAGGAGCAAGAGAGAGCTGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1068
 Db 1267 TGTCTTGAAGGAGCAGCAGAGAGTGTGTAGCGGTTTCCATGACTGCTGCTGCTGCTGCTG 1326
 QY 1069 GACTGCTTCCCGGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1110
 Db 1327 CCTGTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1386
 QY 1111 AATAAGAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1170
 Db 1387 AAGAAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1446
 QY 1171 GAATGAGAGAGGAGCAGCAGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1230
 Db 1447 GCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1306
 QY 1231 ACCCTGAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1290
 Db 1507 CTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1566
 QY 1291 GGGGGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1350

Db 1567 GGGGGGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1626
 QY 1351 GTGTAGTGGGCGCGCCCAAGGCTCTCACGCTGCTGCGCGCGCGCGCTTCCCTC 1410
 Db 1627 GCGCTTCTTTGGGAGACCCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1686
 QY 1411 TGGCTTCAATTTGCAATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1470
 Db 1687 GGTTCACCATCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1746
 QY 1471 AAGATGCCAGCGCTTCCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1530
 Db 1747 AAGTTTCCAGCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1806
 QY 1531 GTCTATGAGCATTTTTCAGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1590
 Db 1807 CTGTTGTGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1866
 QY 1591 CGGCTGAGTCCGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1650
 Db 1867 TGGACCCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1926
 QY 1651 AACCCCACTACCGCAAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1710
 Db 1927 ACAGAGACCAACTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1986
 QY 1711 GTGGTTTCAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1770
 Db 1987 AGTGCCTTTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2046
 QY 1771 TTCTATACCTTCAAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1830
 Db 2047 TGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2106
 QY 1831 TCTGCTTACAGGCGGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1890
 Db 2107 AGCGTCTACAGGCGCAAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2166
 QY 1891 CTGGCCATCAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1950
 Db 2167 AGCAGGCGCTTGGTGGATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2226
 QY 1951 CGCAACAGCGCGCTTCTTCAACAGCATGATCAGGCTTACACCTGAGAGGAGCT 2008
 Db 2227 CTCACAGCAGCAGAGCAGCTTCCAGGCTTCCATTCAGAGACTACAGAGGCGCTGCGCT 2284

RESULT 14
 AX282892
 LOCUS AX282892 2526 bp DNA linear PAT 02-NOV-2001
 DEFINITION Sequence 1 from Patent WO0172842.
 ACCESSION AX282892
 VERSION AX282892.1 GI:16609868
 KEYWORDS
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1
 AUTHORS Walke, D.W., Scoville, J., Donoho, G., Turner, C.A., Friedrich, G.C.,
 Abuln, A., Zamrowicz, B. and Sands, A.T.
 TITLE Novel human 75m proteins and polynucleotides encoding the same
 JOURNAL Patent: WO 0172842-A 1 04-OCT-2001;
 LEXICON Lexicon Genetics Incorporated (US)
 FEATURES
 source 1..2526
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 469 a 770 c 697 g 587 t 3 others
 ORIGIN
 Query Match 19.2%; Score 386.4; DB 6; Length 2526;
 Best Local Similarity 51.1%; Pred. No. 3.7e-54;

Matches 1042; Conservative 1; Mismatches 947; Indels 48; Gaps 4;

```
OY 1 ATCACTTACAGGCGCATCGCATAGCTGCGAGACAGGTGCGCTTCCCGGCTTGTG 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 499 ATTAGTATAGCGCGCCAGCGAGCGAGCGTACGTGAGTACCGAGTATCCCTTTC 558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 61 CGTACCAACCGAGCGCGCCAGCCAGCGTGAAGGCGCATGTGATGTGTCGTC 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 559 CGCACCATCCCAATGACAAATACAGGTGAGAACCATGTGCTGCTGCAAGAGTTC 618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 121 CGCTGGAACATGATCATTTGTGTGTGAGCAGCAGCATTAGGCGCCAGCAATG 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 619 GGGTGAACCTGGATCTCTGTGTGTGAGCAGTACAGTATGAGGAGCAGTGGGGT 678
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 181 CTGCTTGGGAGCGCTGGCGCGCGAGATCTGATGCTTCCAGSAGAGCGTCCC 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 679 GCACGTGAAGAACCGCCACTGTGTAGGGGATCTGATTTCAAGACATCATGCCC 738
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 241 ACACGTGACGCCAACAGAACATGATGATGAGAGCGGCGCGCTGGTGAACATT 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 739 -----TTCTGTGCCAGGTGGCGATGAGAGATGATGATGCTCATG 780
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 301 GACAACTGACAGACAGACAGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTG 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 781 CGCACCCTGGCCAGCGCGCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTG 840
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 361 TACCACTTCTCAATAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 841 AAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 900
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 421 GATCCTGGGCGCATGACCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 480
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 901 GAAAGCTGGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 960
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 481 TTCTGTGGCATACATTCAGAGCGGTGCGCATCCCGGCTGTGTGTGTGTGTGTGTG 540
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 961 GTGCTGTGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1020
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 541 GGGCCACAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 594
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1021 TATGCCGGGCGAGACAGAGAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1080
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 595 AACCAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 654
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1081 AATTCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1140
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 655 TGTGGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 714
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1141 TTCTCATGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1200
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 715 CACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 774
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1201 CACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1260
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 775 CTGCTGTAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 834
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1261 CTTTTGGAGCAGATTCACAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1320
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 835 CCGCAAGGAGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 894
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1321 GACAAAGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1380
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 895 AATCCCTTTCAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 948
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1381 TGGACCTTTCAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1440
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 949 ATCAAGACATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1008
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1441 ACCAAATATCAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1500
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1009 TGGCACTCAGGCGAAAGAAAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1068
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1501 TGTCTTGAAGGCGACAGAGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1560
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
OY 1069 GACTGCTTCCCGGCGACTTCTCAACGACACTGAA-----TGCCTG 1110
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1561 CCTTGGGGGCGTGGGAGCTTCTTCAAGAGAGTCAACCTTACAGATCCAGGCTGTGG 1620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1111 AATTAAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1621 AAGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1680
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1171 GAATGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1681 GCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1740
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1231 ACCCTGCGCATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1290
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1741 CTGCTTGGGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1800
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1291 GGGGGGCGCGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1350
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1801 GGGGGGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1860
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1351 GTGTACGTGGGCGCGCGAGGTCTCACCTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1410
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1861 GGTCTTGTGGGAGAACCAAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1920
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1411 TGTTCACAAATTTGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1470
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1921 GTTTCACCAATCTTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1980
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1471 AAGATGGCGAGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1530
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1981 AAGTTTCCACCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 2040
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1531 GTCTGTAGGCAATTTATACAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1590
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2041 CTGTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 2100
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1591 CGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1650
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2101 TGGACCCGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 2160
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1651 AACCCCACTACCGCAACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1710
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2161 ACAGAGACCAACTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 2220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1711 GTGGGTTTCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1770
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2221 AGTGCCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 2280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1771 TTCAATCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1830
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2281 TGTGTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 2340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1831 TGTGCTTACAGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1890
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2341 AAGCTGTACAGGCGCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 2400
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1891 CTGGGCAATCAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1950
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2401 AGCAGGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 2460
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1951 CGCAACAGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 2008
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2461 CTCAACAGCAGAGCACTTCAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 2518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
RESULT 15
AX282900
LOCUS AX282900 2951 bp DNA linear PAT 02-NOV-2001
DEFINITION Sequence 9 from Patent W00172842.
ACCESSION AX282900
VERSION AX282900.1 GI:16609872
KEYWORDS
```


Db 2572 AGCGTCTACGAGGCGCAGTACCTGCGGCCAACATGATGGCTGGGCTGAGCAGCCTG 2631
OY 1891 CTGGCCATCAGCCTTGCGCTACTTCGGCCCCAAGTCCTACATGATCCTCTTACCCGGAG 1950
Db 2632 AGCAGCGGCTTCGGTGGTATTTCTCGCTAGTGTGATCTCTGCGCCGAGAC 2691
OY 1951 CGCAACAGCGCCGCTACTTCAACAGCATGATCCAGGGCTACCCATGAGAGGAGCT 2008
Db 2692 CTCACAGCAGCAGAGCACTTCAGGCTTCATTCAGAGCTACAGAGGCGGCTGCGGCT 2749

Search completed: May 23, 2003, 16:36:36
Job time : 5982.99 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-p5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 23, 2003, 11:20:10 ; Search time 420.125 Seconds
(without alignments)
10774.220 Million cell updates/sec

Title: US-09-927-315-12

Perfect score: 2010

Sequence: 1 ataccctacgcgcgcacacag.....aacacatgagagagacactag 2010

Scoring table:

IDENTITY-NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N.GeneSeq_101002:*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*

2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*

3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*

4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*

5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*

6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*

7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*

8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*

9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*

10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*

11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*

12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*

13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*

14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*

15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*

16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*

17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*

18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*

19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*

20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*

21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*

22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001.DAT:*

23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*

24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2010	100.0	2010	21	AAZ58965
2	1232.8	61.3	2993	21	AAZ58963
3	1212.6	60.3	2532	21	AAZ58964
4	392.8	19.5	2771	21	AAZ50743
5	386.4	19.2	2292	22	AAAD19502
6	386.4	19.2	2526	22	AAAD19501
7	386.4	19.2	2951	22	AAAD19505
8	385.2	19.2	2759	24	ABL55953
9	376.6	18.7	2333	21	AAZ50745

10	368.8	18.3	2526	22	AAAD17516	Human taste recept
11	359.2	17.9	2579	21	AAZ50744	Mouse sensory tran
12	258	12.8	697	22	AAZ546935	Human G-protein-co
13	258	12.8	697	24	ABK81713	CDNA encoding nove
14	218.6	10.9	2559	24	ABK47345	CDNA sequence enco
15	218.6	10.9	3458	24	AAAL38462	CDNA encoding a hu
16	217	10.8	2559	24	AAZ597395	Human SMC1 gene cD
17	217	10.8	3489	22	AAZ514574	Human CDNA encodin
18	216.2	10.8	2553	22	AAAD17509	Human taste recept
19	187	9.3	4131	18	AAZ69290	Dogfish shark kidn
20	187	9.3	4134	24	ABL59964	DNA encoding shark
21	187	9.3	4134	24	ABK14910	Human novel G-prot
22	184.2	9.2	1101	22	AAAD19503	Human taste recept
23	182.2	9.1	8191	22	AAAD17515	Human G-protein co
24	180.2	9.0	2729	24	ABL55952	Bovine parathyroid
25	172.6	8.6	5275	19	AAZ56962	Bovine parathyroid
26	172.6	8.6	5275	20	AAZ55857	Bovine parathyroid
27	172.6	8.6	5275	20	AAZ55053	Bovine parathyroid
28	172.6	8.6	5275	20	AAZ52483	Bovine parathyroid
29	172.6	8.6	5275	21	AAZ59296	CDNA encoding BOPC
30	172.6	8.6	5275	24	AAZ12120	Human calcium sens
31	167	8.3	2703	21	AAZ50617	Human calcium sens
32	167	8.3	2922	21	AAZ50616	Human calcium sens
33	167	8.3	3003	21	AAZ50615	Human calcium sens
34	167	8.3	3234	20	AAZ50922	Human calcium rece
35	167	8.3	3234	20	AAZ31049	Human CAR coding s
36	167	8.3	3234	21	AAZ51398	Human wld type ca
37	167	8.3	3234	24	AAZ43269	Human CAR cDNA seq
38	167	8.3	3809	19	AAZ61382	Parathyroid calcit
39	167	8.3	3809	19	AAZ6964	Human parathyroid
40	167	8.3	3809	19	AAZ55859	Human parathyroid
41	167	8.3	3809	20	AAZ55055	Human parathyroid
42	167	8.3	3809	20	AAZ52485	Human parathyroid
43	167	8.3	3809	21	AAZ59298	Human calcium rece
44	167	8.3	3809	22	AAZ68094	Small human parath
45	167	8.3	3809	24	AAZ12122	CDNA encoding pHUC

ALIGNMENTS

RESULT 1	AAZ58965	standard; DNA: 2010 BP.
ID	AAZ58965	
AC	AAZ58965	
DT	08-MAY-2000	(first entry)
XX		
DE	Human GPCR-B4 polypeptide encoding DNA.	
XX		
KW	Sensory transduction G-protein coupled receptor; GPCR; GPCR-B4; human;	
KW	taste transduction pathway; taste receptor; foliate; fungiform; food;	
KW	circumvallate; taste signaling; pharmaceutical; ds.	
XX		
OS	Homo sapiens.	
XX		
EH	key	Location/Qualifiers
FT	CDS	1..2010
FT		/tag= a
XX		/product= "GPCR-B4"
XX		
XX	WO200006593-A1.	
XX		
PD	10-FEB-2000.	
XX		
XX	27-JUL-1999;	99WO-US17104.
PF		
XX		
PR	28-JUL-1998;	98US-0095464.
PR	17-DEC-1998;	98US-0112747.
XX		
PA	(RESC) UNIV CALIFORNIA.	
XX		

PI Zuker CS, Adler JE, Lindemeier J;
XX WPL: 2000-195257/17.
DR P-PSDB; AA177558.
XX
PT New isolated sensory transduction G-protein coupled receptor, useful
PT for developing products for use in studying and modulating the taste
PT transduction pathway and for generating taste topographic maps -
XX
PS Claim 5; Page 73-74; 76pp; English.

CC The invention provides nucleic acids encoding rat, mouse and human
CC sensory transduction G-protein coupled receptor (GPCR) polypeptides. The
CC GPCR polypeptides are components of the taste transduction pathway. The
CC nucleic acids can be used to identify taste cells and as tools for the
CC generation of taste topographic maps that elucidate the relationship
CC between the taste cells of the tongue and taste sensory neurons leading
CC to taste centers in the brain. GPCR-B4 is useful as a nucleic acid probe
CC for identifying subpopulations of taste receptor cells such as foliate,
CC fungiform, and circumvallate taste receptor cells. The polypeptides can
CC be used for identifying compounds that modulate sensory signaling in
CC sensory cells. Such modulators of taste transduction are useful for
CC pharmacological and genetic modulation of taste signaling pathways. These
CC modulatory compounds can then be used in the food and pharmaceutical
CC industries to customize taste. The present sequence represents a DNA
CC encoding a human GPCR-B4 polypeptide.

XX
SQ Sequence 2010 BP; 387 A; 691 C; 515 G; 417 T; 0 other;

Query Match 100.0%; Score 2010; DB 21; Length 2010;
Best Local Similarly 100.0%; Pred. No. 0;

Matches 2010; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATCACTACAGCGGCATGAGATGAGTGGAGACAAGTGGCTCCGCGTTGGTG 60
DB 1 ATCACTACAGCGGCATGAGATGAGTGGAGACAAGTGGCTCCGCGTTGGTG 60
QY 61 CGTACACACCCAGCGCGACACACAGTGCAGGCGCATGTGATGCTGCATTTT 120
DB 61 CGTACACACCCAGCGCGACACACAGTGCAGGCGCATGTGATGCTGCATTTT 120
QY 121 CGGTGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
DB 121 CGGTGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 180
QY 181 CTGCTTGGCGAGCGCGTGGCGCGCGGACATGCTGATCGCTTCCAGAGACCTG 240
DB 181 CTGCTTGGCGAGCGCGTGGCGCGCGGACATGCTGATCGCTTCCAGAGACCTG 240
QY 241 ACACGTGAGCCCAACAGAACATGATGATGATGATGATGATGATGATGATG 300
DB 241 ACACGTGAGCCCAACAGAACATGATGATGATGATGATGATGATGATGATG 300
QY 301 GACAACTGAG 360
DB 301 GACAACTGAG 360
QY 361 TACCACTTCTTCAATGAGTGTGCTGCGAGAACTTCAAGGCGCGGCTGGATG 420
DB 361 TACCACTTCTTCAATGAGTGTGCTGCGAGAACTTCAAGGCGCGGCTGGATG 420
QY 421 GAGTCTGAGGCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 421 GAGTCTGAGGCGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 TTCTTGGGCGATGACCATGACAGAGGCTGCCATCCCGGGCTTCAATGATGAG 540
DB 481 TTCTTGGGCGATGACCATGACAGAGGCTGCCATCCCGGGCTTCAATGATGAG 540
QY 541 GGGCCACAGAGGTGGGCGCGCCCTGACAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 541 GGGCCACAGAGGTGGGCGCGCCCTGACAGAGAGAGAGAGAGAGAGAGAGAGAG 600

QY 601 GAGTGCAGCAACTGCTGTAAGCGCCACCTGTGCTTCAACACATTTCAAGGCTCTGG 660
DB 601 GAGTGCAGCAACTGCTGTAAGCGCCACCTGTGCTTCAACACATTTCAAGGCTCTGG 660
QY 661 GAGGCTGTGCTTCAACAGCGTGTACTGCGGTCTATGCTGTGGCCATGCGCTGACAGC 720
DB 661 GAGGCTGTGCTTCAACAGCGTGTACTGCGGTCTATGCTGTGGCCATGCGCTGACAGC 720
QY 721 CTCTGCGGCTGTGCAAAAGACACTGACCAAGAGGGGTGCTTACCCCTGGACGCTGCT 780
DB 721 CTCTGCGGCTGTGCAAAAGACACTGACCAAGAGGGGTGCTTACCCCTGGACGCTGCT 780
QY 781 GAGAGATCTGGAAGTCAACTTCACTCTCTGAGCACCACCAATCTTTCGACCCCA 840
DB 781 GAGAGATCTGGAAGTCAACTTCACTCTCTGAGCACCACCAATCTTTCGACCCCA 840
QY 841 GGGAGCTGGCTGTGACACTTGGAGATTGTCCAGTGGCAATGGAGCGGACAGAAATCC 900
DB 841 GGGAGCTGGCTGTGACACTTGGAGATTGTCCAGTGGCAATGGAGCGGACAGAAATCC 900
QY 901 TTCCAGAGCGTGGCGCTCTACTACCTACCCCTGAGGAGAGAGAGAGAGAGAGAGAG 960
DB 901 TTCCAGAGCGTGGCGCTCTACTACCTACCCCTGAGGAGAGAGAGAGAGAGAGAGAG 960
QY 961 CTGCACACCGTCAACACAGATCCCTATGCTCATGTGTTCAAGAGGTGCCAGTCAAGG 1020
DB 961 CTGCACACCGTCAACACAGATCCCTATGCTCATGTGTTCAAGAGGTGCCAGTCAAGG 1020
QY 1021 CAAAAGAAAGAGCTGTGGGATCAGCTGCTGCTTCAGATGATGATGATGATGATG 1080
DB 1021 CAAAAGAAAGAGCTGTGGGATCAGCTGCTGCTTCAGATGATGATGATGATGATG 1080
QY 1081 GGCACCTTCTCAACACAGATGATGATGATGATGATGATGATGATGATGATGATG 1140
DB 1081 GGCACCTTCTCAACACAGATGATGATGATGATGATGATGATGATGATGATGATG 1140
QY 1141 TCTGCTTCAAGGCGGAGCTGCTTCTGTAATGATGATGATGATGATGATGATGATG 1200
DB 1141 TCTGCTTCAAGGCGGAGCTGCTTCTGTAATGATGATGATGATGATGATGATGATG 1200
QY 1201 GCGCTGTGGCGCGCGCTGGGCTTCTGAGACCTGCGCATGCTGGATGATGATGATG 1260
DB 1201 GCGCTGTGGCGCGCGCTGGGCTTCTGAGACCTGCGCATGCTGGATGATGATGATG 1260
QY 1261 CACTTCCAGACACCATAGTGTGCTGCTGAGGCGGCGGCTGCTGATGATGATGATG 1320
DB 1261 CACTTCCAGACACCATAGTGTGCTGCTGAGGCGGCGGCTGCTGATGATGATGATG 1320
QY 1321 CTGCTGTGGTGGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
DB 1321 CTGCTGTGGTGGCAATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
QY 1381 TGGCTTGGCGGAGAGCGGCTTTCCTGCTGCTGATGATGATGATGATGATGATG 1440
DB 1381 TGGCTTGGCGGAGAGCGGCTTTCCTGCTGCTGATGATGATGATGATGATGATG 1440
QY 1441 GTGGGCTTTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
DB 1441 GTGGGCTTTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1500
QY 1501 AGCTACTGGGTGCGCTACAGAGGGGCGCTACGCTCTATGATGATGATGATGATG 1560
DB 1501 AGCTACTGGGTGCGCTACAGAGGGGCGCTACGCTCTATGATGATGATGATGATG 1560
QY 1561 ATGCTATTTGTTGTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
DB 1561 ATGCTATTTGTTGTAATGATGATGATGATGATGATGATGATGATGATGATGATG 1620
QY 1621 GATGACCCCAAGATCAATTTGCTCTGTAACCCCAAGTACAGGCAAGCAGCAGCTG 1680
DB 1621 GATGACCCCAAGATCAATTTGCTCTGTAACCCCAAGTACAGGCAAGCAGCAGCTG 1680
QY 1681 AACACAGCCTGAGACCTGCTCTGATGATGATGATGATGATGATGATGATGATG 1740

Db 1345 CAAAGGGACATGCGATGCTTGTGACATCAATCCAGTGGAGCTGAGCCAGAT 1404
 Oy 898 CCCCTCCAGAGGCTGGCTCTCTACTACGCCCCCTGAGGACGACGAGAAAT---CAG 954
 Db 1405 CCCCTCCAGAGGCTGGCTCTCTACTATCTCCACGACGACGAGGCTTACTTACTTAAAT 1464
 Oy 955 ACATCTCTGACACCGCTCAACACAGATCCCTATGTCATGTGTCCAGAGGTGGCAG 1014
 Db 1465 GTGTCTGCTGATACCCCAACACAGGCTGCTCTCATGTGTCCAGAGGCTGGCAG 1524
 Oy 1015 TCAGGGCAAAAGAAAGGCTGTGGGCTTCAGCTCTGCTCTGAGTGCATGACTGC 1074
 Db 1525 CCAGGGCAAAATGAAAGTCTGTGGGCTCCACCTTGTCTTCGAGTGCCTGGATTGT 1584
 Oy 1075 CTTCGGGACCTCTCTCAACACAGTGA-----TGCCCGAAT 1113
 Db 1585 ATGCGAGGACCTTACCTCAACGCTCAAGAGATGTTAACTGTCTGCTCCGCGGT 1644
 Oy 1114 AACGAGTGTCTTACAGAGTGAAGCTCTGCTCAAGCGGAGCTGTCTCTGGA 1173
 Db 1645 TCCATGTGTCTTACAGAGATGATCATCTGCTTCCAGGGGCGGCTTCTCTGAG 1704
 Oy 1174 TGGCATGAGGACCCACCATATGCTGTGGCCCTGCTGGCCCTTGGGCTTCTTACGAC 1233
 Db 1705 TGGGACAGAGTGGCCACCATGCTGTGGCCCTGCTGTGGGCTTCTTACGATCA 1764
 Oy 1234 CTGGCCATCTGCTGATATTTGTGAGGACATCTTCAGACACCATATGTTGCTGGCTGG 1293
 Db 1765 CTGGCCATCTTCTTCTTCTTCTTCTGAGACATTTCCAGACACCATATGCTGGCCG 1824
 Oy 1294 GGGCCCATGTCTCTGATGCTGACACATGCTGTGTGGGATACATGCTGCTGG 1353
 Db 1825 GGGCCCATGTCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1884
 Oy 1354 TAGCTGGGGGGGGGAGGCTTCCAGCTGCTGCTGGCGGCGGCTTCTTCTGCTGCTG 1413
 Db 1885 TATGTGGGGGGGGGAGGCTTCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1944
 Oy 1414 TTCACAATTTGCAATCTCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1473
 Db 1945 TTCTCAATCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2004
 Oy 1474 ATGGCCAGCGGCTTCCAGCGGCTTACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1533
 Db 2005 ATGGCCAGCGGCTTCCAGCGGCTTACAGTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2064
 Oy 1534 TCTATGCAATTTACAGGCTTCAAAATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1593
 Db 2065 TTCTGCTGCTTCAATGAGGCTTCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2124
 Oy 1594 CTTGAGTCCGACCC---CCGTAAGTACCCGATGACCCCAATGATGCTGCTGCTGCTG 1650
 Db 2125 ACATCAACCCGATGAGGCTGAGGACCGGATGACCCCAATGATGCTGCTGCTGCTGCTG 2184
 Oy 1651 AACCCCACTACGCAAGAGCTGCTGCTTCAACACAGCTGCTGCTGCTGCTGCTGCTGCTG 1710
 Db 2185 CACCTCACTACGCAAGGCTGCTGCTTCAACACAGCTGCTGCTGCTGCTGCTGCTGCTG 2244
 Oy 1711 GTGGGTTTACGCTTGGCTTCAATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1770
 Db 2245 CTGGGTTTACGCTTGGCTTCAATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2304
 Oy 1771 TTCTCAACCCGATGAGGCTTCAATGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1830
 Db 2305 TTCTCACTGCTCAAGATGAGCTTCTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2364
 Oy 1831 TCTGCTCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1890
 Db 2365 TCTGCTCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2424
 Oy 1891 CTGGCCATGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1950

Db 2425 CTGGCCATGCGCTTGGGATCTTGTGCCCCCAAGTGTATACATGATCTTTCTACCCGGAG 2484
 Oy 1951 CGAACACAGCGCGCTACTTCAACAGATGATCCAGGGCTACCATGAGGAGGACTAG 2010
 Db 2485 CGAACACAGCGCGCTACTTCAATGATGATCCAGGGCTACCATGAGGAGGACTAG 2544
 RESULT 3
 AA258964
 ID AA258964 standard; DNA: 2532 BP.
 XX
 AC AA258964;
 DT 08-MAY-2000 (first entry)
 XX
 DE Mouse GPCR-B4 polypeptide encoding DNA.
 KW Sensory transduction G-protein coupled receptor; GPCR; GPCR-B4; mouse;
 KW taste transduction pathway; taste receptor; foliate; fungiform; food;
 KW circumvallate; taste signaling; pharmaceutical; ds.
 XX
 OS Mus sp.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..2532
 FT /*tag= a
 FT /product= "GPCR-B4"
 PN W0200006593-A1.
 PD 10-FEB-2000.
 XX
 XX 27-JUL-1999; 99MO-US17104.
 PF
 XX 28-JUL-1998; 98US-0095464.
 PR 17-DEC-1998; 98US-0112747.
 XX
 PA (REGC) UNIV CALIFORNIA.
 PI Zuker CS, Adler JE, Lindemeyer J;
 XX WPI: 2000-195257/17.
 DR P-PDB: AA177357.
 XX
 PT New isolated sensory transduction G-protein coupled receptor, useful
 PT for developing products for use in studying and modulating the taste
 PT transduction pathway and for generating taste topographic maps -
 XX
 PS Claim 5; Page 71-72; 76pp; English.
 XX
 CC The invention provides nucleic acids encoding rat, mouse and human
 CC sensory transduction G-protein coupled receptor (GPCR) polypeptides. The
 CC GPCR polypeptides are components of the taste transduction pathway. The
 CC nucleic acids can be used to identify taste cells and as tools for the
 CC generation of taste topographic maps that elucidate the relationship
 CC between the taste cells of the tongue and taste sensory neurons leading
 CC to taste centers in the brain. GPCR-B4 is useful as a nucleic acid probe
 CC for identifying subpopulations of taste receptor cells such as foliate,
 CC fungiform, and circumvallate taste receptor cells. The polypeptides can
 CC be used for identifying compounds that modulate sensory signaling in
 CC sensory cells. Such modulators of taste transduction are useful for
 CC pharmacological and genetic modulation of taste signaling pathways. These
 CC modulatory compounds can then be used in the food and pharmaceutical
 CC industries to customize taste. The present sequence represents a DNA
 CC encoding a mouse GPCR-B4 polypeptide.
 XX
 SQ Sequence 2532 BP; 498 A; 824 C; 631 G; 579 T; 0 other;
 Query Match 60.3%; Score 1212.6; DB 21; Length 2532;
 Best Local Similarity 76.3%; Pred. No. 9.3e-245;
 Matches 1555; Conservative 0; Mismatches 454; Indels 30; Gaps 4;
 Oy 2 TCACCTACAGCGCCATCAGGATGAGCTGGAGACAAGTGGCTTCCGGCTTGTGCTGC 61

Db	494	TCACATATAGGGGCATACCGACAAAGCTGCAAGACAGGCGGCTTCCTCGCATGCTGC	553
QY	62	GTACCAACACCAGGCGCGACACACAGGTGAGGCGCATGTGTGAGTGAATGTGCATCTCC	121
Db	554	GCACTGTGCCAAGGCGACCCACACACATGAGGCCATGTGTGCACTGATGTCTACTCTCC	613
QY	122	GCTGGAACTGGATCANTTGTGCTGTGTAGCAGCAGCACTATGSGCGCGACAAATGGCCAGC	181
Db	614	AGTGGAACTGGATCTGCTGTGTGTGTAGCGATGTACATTTATGTGGCGAGAAACACACC	673
QY	182	TGCTTGGGCGAGCGGTGGC---CCGGCGGACATCTGCATCGCCCTTCCAGAGACGCTGC	238
Db	674	TGCTGAGCCACGCGCTGCACCAACATGGGGACATCTGCATTCGCTTCCTCCAGAGGTTCTGC	733
QY	239	CCATACATGCAGCCCAACACAGAACATGACCTCAGAGAGGGCGGCGCCCTGGGACATTTG	298
Db	734	CCGTACCAAGAACCCAAACACAGGCTGTGTAGGCTGTGAGGACGAGAACCACTGGACAACTCC	793
QY	299	TGACAAAGCTGCACAGCAGACACAGCGCGCGGTGTGTGTGTCTGTCTTCGCGCGACCTGACC	358
Db	794	TGACAAAGCTGCAGCGGAGCTTGGCGCGGTGTGTGTGTATTTCTTCGCGCGAGCTGACC	853
QY	359	TGTACCACTTCTTCATGTAGGTCTCTGCCAGAACTTCAACGGGCGCGTGTGTGATCGCT	418
Db	854	TGCACAACTTCTTCGAGAGTCTCGGTGAACTTCAACGGGCTTGTGTGTGATTCCT	913
QY	419	CCGAGTCTCTGGGGCATTCAGACCCGCGCTCTGCACAACTCCAGGAGTGGGGCACTTGGGCA	478
Db	914	CTGAGTCTTGGGCGCATTCAGACCCGTGTCTACACAACTTCAACAGTGTGGCCACACGGGCA	973
QY	479	CTTCTCTGGGCGATTCACCATTCAGAGCGTGCCCATCCCGGGCTTTCAGTAGTTCGCGAGT	538
Db	974	CTTCTCTGGGCTGTACCATTCAGAGGGTTCATCCCTGGTTCAGCCAGTTCCGAGTGC	1033
QY	539	GGGGCCCAAGGCTGGGCGCGCCACCCCTGACAGAGCAACGCCAGACTTATCTGCAACC	598
Db	1034	GCCATGACAAAGCAGGGTATTCGATTCCTTAAAGAGACACGCGTGGGCACTACTGTAAAC	1093
QY	599	AGGATGTGCGACACTGGCTCTTAACGGCACCTGTGCTTCAACACAAATTCACAGGCTCTG	658
Db	1094	AGGACTGTGACGCTGTGATATCATATCACTAGTAGTCCCTTCAACAGGTTCTTCAAGCTTGTGG	1153
QY	659	GGGAGCGTGTCTGTACAGCGTGTACTTGCGGTATATCTGTGGCCATGTGCCCTGTGCACA	718
Db	1154	GGGAGCGTGTGTGTACAGCGTGTACTTGCGGTGTATCTGTGGCCATGTGCCCTGTGCACA	1213
QY	719	GCTCTCTCGGCTGTGACAAAAGCACTGCACCAAGAAGGTGTCTTACCCCTGTGGACGTGC	778
Db	1214	GACTCTCTCACTGCATCAACAGGTCCGCTGTGCACCAAGAAATCTGTATTCATATGTGGACGTAC	1273
QY	779	TTTGAAGGATCTGGAAGGTCAACTTACATCTCTGGACCAACAAATTCCTTTCAGACCGCG	838
Db	1274	TCAGAGGATCTGTGCATGTATTAACCTTACGCTCTGGGCAACCAAGCTCTTCTTTCAGCAAC	1333
QY	839	AAGGGAGCGTGGCTGTGTCACTTGGAGATTTGTCAGTGGCAATAGGGACCGGACCGAATTC	898
Db	1334	AAGGGAGCGATGGCGATGTCTGTGTGACATATCAAGTGGAGTGGGGCTGTGAGCCGAAAC	1393
QY	899	CTTTCACAGAGGTGCGCTCTTACTTACCCCTGACGCGACGCTGAGAAACAT--CAAGA	955
Db	1394	CTTTCACAAAGATGCGCTCTTACTTCCCCACGAGACGAGCGAGGTGACCTTATAGCAATG	1453
QY	956	CATCTGTGCACACCGTCACAAACAGATATCCATATGTCCATGTGTTCCAAAGAGTGGCCAGT	1015
Db	1454	TGTCTGTGTACACCCCAACAAACAGGTCCCATATCACTGTGTCTTAAAGTGTTCACGC	1513
QY	1016	CAGGGCAAAAGAAAGGCTGTGGGATTCACAGTGTGCTGTTCAGTGCATTCAGATGCC	1075
Db	1514	CTGGGCAAAATGAAAACCATATGGGCTTCACCCATGTCTGTGAGTGTGTGACATCTC	1573
QY	1076	TTCCCGGACACTTCTTCAACACACTGAA-----TGCCGGAATA	1114

Db	1574	CGCGGACACCTACCTCAACCGATCAGTAGATAGTTTAACTGTCTGTCCGCCGGGTT	1633
Qy	1115	ACGAGTGGTCCCTACAGATGAGACCTCTGCTTCAAGGGGAGCGTCTCTGAGAT	1174
Db	1634	CGATGTGGTCTTTACAAGAACAAACATGCCCTTGCTTCAAGGGGGGCGTGGCTTCCGAGAT	1699
Qy	1175	GGCATGAGGACACCAACCATCGCTGTGTGGCCCTGCTGGCCGCGCTTCTTAGACACC	1233
Db	1694	GGCACGAGTGGCCACATATGATGGTGAACATCTGGCCGCGCTTCAATCAGTACGC	1755
Qy	1235	TGGCCATCTGTGATATATTCGGAGGGACATCCAGAACCCATAGTTCGCTGGGCTGGGG	1299
Db	1754	TGGCCATCTGTCTCATCTTCTTGGAGACATTTCCAGACCGCCCATGCTGGCGGGGG	1813
Qy	1295	GCCCCATGTGCTTCATGATGTCACATGCTGCTGGTGGCATCATAGTGTCCGGTGT	1354
Db	1814	GCCCCATGTGCTTCTCATAGTGTGGTGGCCCTGCTGCTGGGCTTCCGGATGGTCCCGTGT	1877
Qy	1355	ACGTGGGGCGCCCAAGGTCTCAACCTGTGCTGGCCGACAGGCCCTCTTCCCTGTGCT	1414
Db	1874	ATGTGGGGCCCCCAGAGTCTTCTCTGTTCTGGCGGACAGGCTTCTTCAACGGTTGCT	1933
Qy	1415	TCACAAATTGCATCTCTGTATATGCGCGTGGGCTTCTTCCAGATCTGTGCGCTTCCAGA	1477
Db	1934	TCTCGTCTGTCTCTCTCTGTGATCAACGGTGGGCTCTTCCAGATTTGTGTCTGTCCAGA	1995
Qy	1475	TGGCCAGCGGCTTCCAGCGGCGCTACAGCTACAGTGGGGTCCGCTACGAGGGGCGCTACGCTGT	1533
Db	1994	TGGCCAGCGGCTCTCCAGAGCGCTTACCGGTTTCTGTGATGCTTACACGGGCGCTTACGCTGT	2055
Qy	1535	CTATGGCATTTATACAGGTACTCAAAATGGTCTATTGTGTAATTTGGCATCTGGACAGCG	1594
Db	2054	TGTGTGGCTTCATCAACGGGCGGTCAAAAGTGGCCCTGGTGGCGGACCAATGTGGCGACCA	2113
Qy	1555	CTCACTCCACACC--CGTATCTAGTACCCTGATGACCCCAAGATCACAATTTGTCTGTGA	1651
Db	2114	CCATCAACCCCATTTGGCGGGACCGACCCCGATGACCCCATATATCAATATCTCTCTGGCC	2172
Qy	1652	ACCCCAATCTACCGGAACAGCGCTGTGTTCAACACAGCCGTGGAGCTGTGCTCTAGTGG	1711
Db	2174	ACCCCAATCTACCGGAACAGCGGCTACTCTTCAACACAGCATGATGATCTGTGTCCGTGC	2233
Qy	1712	TGGGTTTCACTTGCCTTACATGGGGCAAAAGCTGGCCACCACTACCAAGAGGCCAAGT	1777
Db	2234	TGGGTTTCACTTGCCTTACAGTGGGGCAAAAGACTGGCCACCACTACCAAGAGGCCAAGT	2299
Qy	1772	TCATCACTCTAGCATGACATCTCTATTTTCACTCATATCCGTCTCCCTGTGACCTTCATGT	1833
Db	2294	TCATCACTCTAGCATGACATCTCTCTCTCTCACTCTCTCACTCTCTCTGACAGTTTCACT	2355
Qy	1832	CTGGCTTACAGGGGGTCTGGTCAACATGGTGGAGCCCTTGGGTCATGTGTCACAACTCC	1899
Db	2354	CTGTCACTAGATGGGTGTGTGTCACTATGATTTCTCTGTCACTGTGTCTCAACTTTC	2413
Qy	1892	TGGCCATCAGCCTGGGTACTTCTGGCCCCCAAGTGTACATGATCTCTTCAACCGGAGC	1951
Db	2414	TGGCCATCAGCCTGGGTGTGTGTGTCACTATGATTTCTCTGTCACTGTGTCTCAACTTTC	2477
Qy	1952	GCACACGCCCGCTACTTCAACAGCATGTATCCAGGGCTTACACATGAGAGGAGACTAG	2010
Db	2474	GCACACACTTCACTTATTTTCAATAGCATGTATTCAGGGCTTACACATGAGAGGAGACTAG	2532
RESULT 4			
AAZ50743			
ID	AAZ50743 standard; cDNA; 2771 BP.		
AC	AAZ50743:		
XX	31-MAY-2000 (first entry)		
DE	Rat sensory transduction G-protein coupled receptor-B3 cDNA.		

LT 4
0743
AAZ50743 standard; cDNA: 2771 BP.
AAZ50743;
31-MAY-2000 (first entry)
rat sensory transduction G-protein coupled receptor-B3 cDNA

QY 1352 TGAGCTGGGGGCGCCAGAGTCTCCACTGCTGCGCGGAGCCCTTTCCCTCT 1411
 DB 1915 GCTTCTGGGAGGCCAGGTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1974
 QY 1412 GCTTACAAATTTGCAATCTCTGATGCGGCTGCTGCTGCTGCTGCTGCTGCTG 1471
 DB 1975 GGTTCGCAATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2034
 QY 1472 ACATGGCCAGCCGCTTCCAGCGGCTTACAGCTACTGCTGCTGCTGCTGCTGCTG 1531
 DB 2035 AGTTTCTTCAAGAGGAGCCCAATCTTACCTGCTGCTGCTGCTGCTGCTGCTGCTG 2094
 QY 1532 TCTCTATGGCAATTTATCAGGCTACCAAAATGCTATGCTGCTGCTGCTGCTGCTG 1591
 DB 2095 TATTCCTCATTTGCTACGCTCCAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2154
 QY 1592 GGCCTCAGTCCAGCCCGCTTACCTGACCCGATGACCCCAAGATCAATGCTGCTGCTG 1651
 DB 2155 GGACCCAGAGCCAGCCAGGAAATCCAGCGCTTCCCGATGCTGCTGCTGCTGCTGCTG 2214
 QY 1652 ACCCCACTACCGCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1711
 DB 2215 CAGAGGTCAACTCTGAGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2274
 QY 1712 TGGGTTTCACTTCCGCTTACATGGGCAAGAGTGGCCCAACATCAACAGAGGCCAAGT 1771
 DB 2275 GTACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2334
 QY 1772 TCATCACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1831
 DB 2335 GTGTCACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2394
 QY 1832 CTGCTCAAGCGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1891
 DB 2395 GCATTTTCAAGCGGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2454
 QY 1892 TGGCCATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1951
 DB 2455 GGGGGGCTTCAAGCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2514
 QY 1952 GCACACGCGCGCTTCAACAGCATGATCCAGGGCTTACACCATGAGAGGGAC 2007
 DB 2515 TCACAAATACAGAACACTTTCAGGCTCCATCCAGAGCTACAGAGGCGCTGCGGG 2570

 RESULT 5
 AAD19502
 ID AAD19502 standard; cDNA; 2292 BP.
 AC AAD19502;
 XX
 DT 18-DEC-2001 (first entry)
 DE Human novel G-protein coupled receptor (NGPCR) cDNA #2.
 XX
 KW Human; G-protein coupled receptor; GPCR; gene therapy; drug screening;
 KW antisense-therapy; signal transduction; behavioural disorder; obesity;
 KW heartbeats rate; inflammation; immune disorder; diabetes; cancer;
 KW coronary disease; ss.
 OS Homo sapiens.
 XX
 XX Key Location/Qualifiers
 FT CDS 1..2292
 FT /tag- a
 FT /product- "Human NGPCR protein"
 FT /tag- b
 FT /note- "Encodes Ser"
 FT 880..882
 FT /tag- c
 FT /note- "Encodes Ala"
 XX

PN W0200172842-A2.
 XX 04-OCT-2001.
 PD 28-MAR-2001; 2001WO-US09996.
 XX 28-MAR-2000; 2000US-192978P.
 PF (LEXI-) LEXICON GENETICS INC.
 PR Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abulin A;
 XX Zambronicz B, Sands AT;
 XX WPI: 2001-616474/71.
 DR P-PSDB; AAE11970.
 XX
 PT Novel isolated polynucleotides encoding human G protein coupled
 PT receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
 PT expressed NGPCRs for diagnosis of disease, and as probes or primers
 PS Claim 4; Page 74-75; 80pp; English.
 XX
 CC The present sequence is a cDNA encoding human novel G-protein coupled
 CC receptor (NGPCR) protein. NGPCRs are transmembrane proteins that span
 CC the cellular membrane and are involved in signal transduction after
 CC ligand binding. The NGPCR polynucleotide sequences, are useful
 CC in diagnosis and treatment of a disease involving NGPCR, for detecting
 CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
 CC disease, for screening drugs effective in treatment of symptomatic or
 CC phenotypic manifestations of perturbing the normal function of NGPCR in
 CC the body or abnormalities in the signal transduction pathway mediated by
 CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
 CC trial monitoring and/or the treatment of physiological (heartbeats rate)
 CC or behavioural disorders. NGPCR is useful for identifying compounds
 CC useful in the therapeutic treatment of obesity, inflammation, immune
 CC disorders, diabetes, heart and coronary disease, metabolic disorders and
 CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
 CC genomic library which is helpful for identifying polymorphisms,
 CC determining the genomic structure of a given locus/allele and designing
 CC diagnostic tests.
 SQ Sequence 2292 BP; 437 A; 690 C; 638 G; 524 T; 3 other;

 Query Match 19.2%; Score 386.4; DB 22; Length 2292;
 Best Local Similarity 51.1%; Pred. No. 1.5e-71;
 Matches 1042; Conservative 1; Mismatches 947; Indels 48; Gaps 4;

 QY 1 ATCACCCTACAGCGCCATCAGCGATGAGTCGAGACAGAGTGGGCTTCCGCTTGGCTG 60
 DB 265 ATTAGCTATGCGGCGCCAGAGCGAGCGTACAGCGGAGATATCCCTTTCCG 324
 QY 61 CGTACCACACCCAGCGCGCCAGCACAGCTGAGAGCCATGCTGATGCTGCACTTC 120
 DB 325 CGCACCATCCCAATGACAGTACAGTGGAGACCATGCTGCTGCTGCTGCTGCTG 384
 QY 121 CGCTGGAATGATCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
 DB 385 GGGTGGACTGTGATCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 444
 QY 181 CTGCTTGGGAGCGCGCTGCGCGCGAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
 DB 445 GCACGTGAGAACAGAGCCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 504
 QY 241 ACACCTGACACCCACAGACATGACGTGAGAGAGCCAGCGCTGCTGCTGCTGCTGCTGCTGCTG 300
 DB 505 -----TTCTCTGCGCAGAGTGGCGAGAGAGTGAAGTGAAGTCCCTCAGT 546
 QY 301 GACAAGCTGACAGAGAGAGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
 DB 547 CGCACCTGCGCCAGCGCGCGCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 606
 QY 361 TACCACTTCTCAATGAGGTGCTGCGCAGAACTTACAGGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420

Db 607 AGGTTGTTTTGAGTCGCGTGGTGTGACCAACTGACTGGCAGAGTGTGGCTCCCTCA 666
 QY 421 GAGTCTGGGCGCATGACCCGGTCTGGACAACTGACGAGAGCTGGCCACTTGGGCAAC 480
 Db 667 GAGGCTGGGCGCTCTCCAGGACATCACTGGGTTGCCGGATCCAGGCACTTGGGATG 726
 QY 481 TTCTGGGCAATCAACATCCAGAGCGTCCCATCCGCGCTTCAGTAGAGTTCGCGAGTGG 540
 Db 727 GTGCTGGGCGCTGGCATCAGAGAGAGGCTGTCTGCGCTGAGAGCGCTTGAAGAGCC 786
 QY 541 GGGCCAGAGGCTGGGCGCCGACCCCTCAGCAGAGACAGCCAGAGTAT-----ACCTGC 594
 Db 787 TATGCCCCGGGAGACAGAGAGGCCCCCTAGGCGCTTCACAGAGGCTCTGTGTGAGAGC 846
 QY 595 AACCGAGAGTGGACAACTGCTTAACGCCACCTGTCTTCACAACTTCAGAGCTC 654
 Db 847 AATGAGCTCTGAGAGATGCCAACTTTCATGACACAGATCCCAAGTCAAAAGC 906
 QY 655 TCTGGGAGCGTGTCTGACAGCGTGTACTCTGCGGTCTATGCTGTGGCCCATGCCCTG 714
 Db 907 TTCTCCATGAGTTCCTGCTACAGCAGATACGCGGCTGTGTATGCGTGGCCCATGCCCTC 966
 QY 715 CACAGCTCTCGGCTGTGACAAAAGCACTGACCAAGAGGAGTGTACCCCTGGAG 774
 Db 967 CACAGCTCTCGGCTGTGCTGCTGAGCTGTCTCAGGAGGCGGAGCTCACCCTGGAG 1026
 QY 775 CTGCTTGAAGAGATCTGAGAGGTCACTTCACTCTCTGACACCAAACTCTTCTGAC 834
 Db 1027 CTTTGGAGAGATCCACAGAGTGTCTTCTTACACAGAGACAGCTGTGGCTTTAT 1086
 QY 835 CCGCAGAGGAGCTGCTGTGACCTTGAGATTTGCAAGTGGCAATGGAGCCGAGCCAG 894
 Db 1087 GACAAACAGAGATCCCTCAGTATGATTAATGCTGCGAGCTGAGATGAGACCCAG 1146
 QY 895 AATCCCTCCAGAGGCTGGCCCTCTAC-----TACCCCTGACAGAGAGCTGAAGAC 948
 Db 1147 TGGACCTTCAGGCTCTGCTGCTCTCCACAGTGTCTCCAGTCTAACTAAATATAG 1206
 QY 949 ATCAAGACATCTCTGACACAGCTCAACAAAGATCCCTATGTCATGTGTCAAGAG 1008
 Db 1207 ACCAAATTCAGTGGCAGAGGAAAGACAAAGCTGCTAAGTGTGTCTCCAGCAG 1266
 QY 1009 TGGCAGTGAAGGCAAAAGAGAGAGGCTGTGGGATCCAGCTGTGCTTGCATGTCATC 1068
 Db 1267 TGTCTGAAGGCGACACAGAGAGTGTGAGGTTTCATCACTGCTGTTGAGTGTG 1326
 QY 1069 GACTGCTCCCGGACCTTCTCAACCACTGAA-----TGGCCG 1110
 Db 1327 CCTGTGGGCGTGGAGCTTCTCAACAGAGTGACTCTTCAGATGCCAGGCTGTGGG 1386
 QY 1111 AATAAGAGTGTCTCTACAGAGTGAAGCTCTCTGCTTCAAGCGGAGCTGTCTTCTG 1170
 Db 1387 AAGAGAGTGGGCACTGAGGAGAGCCAGACCTGCTTCCGCGGACAGTGTGTTTGG 1446
 QY 1171 GATGAGCATGAGGACCCCACTGCTGTGGCCCTGTGCGCCCTGTGGCTTCTCTACAC 1230
 Db 1447 GCTTGTGCTGAGACACCTTGTGGTGTGCTGAGAGCTTACACCTGCTGCTGCTG 1506
 QY 1231 ACCCTGGACATCTGCTGATATCTGAGAGCACTTCCAGACACCAATGAGTGGCGCT 1290
 Db 1507 CTGCTTGGGAGCTGCTGCTGCTGTGGGCACTGAGACACCCCTGTGTGAGGTACGA 1566
 QY 1291 GGGGCGCCCATGCTGCTCTGATGCTGACAGCTGCTGTGAGGCAATGATGTTGCTCCG 1350
 Db 1567 GGGGCGCCCGCTGTCTTCTTATGCTGGGCTCCCTGGAGAGAGTGGAGGCTCTAT 1626
 QY 1351 GTGACGTGGGCGCCCAAGGTCTCACTGCTGTGCGCGGAGGCGCTCTTCCCTC 1410
 Db 1627 GGGCTCTTGGGAGAACCAAGGCTGCTGCTGTGCTGACGAGGCGCTCTTCCCT 1686
 QY 1411 TGCTTCACAATTTGATCTCTGATGCGCGCTGTCTTTCACATGCTGTGCGGCTTC 1470
 Db 1687 GGTTCACCACTTCTCTGCTGCTGCTGAGAGTGCCTCATTTCCAACTAATCATATCTTC 1746

QY 1471 AAGATGGCCAGCCGCTTCCACGCGCTACAGACTAGTGTCCGCTACAGAGGCGCTTAC 1530
 Db 1747 AAGTTTCCACCAAGTACTACATTTTACACAGCGCTGGGTCCAAACACAGGCTGTGCG 1806
 QY 1531 GTCTTATGACATTTATACGAGTACTCAAAATGCTATGTGTGATTTGATGTCATGCGCA 1590
 Db 1807 CTGTTTGTGATGATGACGTCAAGCGCCAGCTGCTTATCTGTCACTTGTGCTGTG 1866
 QY 1591 CGGCTCAGTCCACCCCGGACAGTACCCGATGACCCCAAGATACATTTGTCTCGT 1650
 Db 1867 TGGACCCCACTGCTGTGAGGAAATACAGCGCTTCCCAATGTGTGATCTGTGAGTGC 1926
 QY 1651 AACCCCACTACCCGACAGCGCTGTCTTCAACACAGCTGAGACTGTCTGTCAAGTGT 1710
 Db 1927 ACAAGACCAACTCCCTGGGCTTCACTAGGCTTCTCTACAAATGGCTCTCTCCATC 1986
 QY 1711 GTGGTTTACGCTTGCTTACATGGGCAAGAGCTGCCCCAACATACAGAGGCCAAG 1770
 Db 1987 AGTGGCTTGGCTGTGAGCTTACCTGTGTAAGACTTGTCCAGAGAACTACAGAGGCCAA 2046
 QY 1771 TTTATCACCTTCAGAGTACCTTCTATTTCACTTCACTGATCCGCTCCCTGTGACCTTCATG 1830
 Db 2047 TGTGACCTTCAAGCTCTCTCTTCACTTCTGCTGTGATGCTGTTCACACAGGCC 2106
 QY 1831 TCTGCTTACAGCGGCTGCTGTGACCACTGTCACCTTGTGTACTGTCTCAACCTC 1890
 Db 2107 AGGCTTACGAGCGGCAAGTACCTGCTGCGGCAACATGATGGCTGAGGAGCCTG 2166
 QY 1891 CTGGCCATCAGCGCTGCTTACTTGGCCCCCAAGTCTTACATGATCTTCTTACCCGAG 1950
 Db 2167 AGCAGCGGCTTCTGGGTATTTTCTGCTTAAAGTCTTCAAGTCTTCTGCGGCCAAC 2226
 QY 1951 CGCAACAGCGCGCTTACTTCAACAGATGATCCAGGCTCACACATGAGAGGAGT 2008
 Db 2227 CTCAACAGACAGAGACTTCCAGGCTTCACTTCAAGGCTTCAAGGAGGAGGCTGCGCT 2284

RESULT 6
 AAD19501
 ID AAD19501 standard; cDNA: 2526 BP.
 AC AAD19501;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Human novel G-protein coupled receptor (NGPCR) cDNA #1.
 XX
 KW Human; G-protein coupled receptor; GPCR; gene therapy; drug screening;
 KW antisense-therapy; signal transduction; behavioural disorder; obesity;
 KW heartbeat rate; inflammation; immune disorder; diabetes; cancer;
 KW coronary disease; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 1..2526
 FT CDS
 FT unsure
 FT /product- "Human NGPCR protein"
 FT 319..321
 FT /tag- b
 FT /note- "Encodes Ser"
 FT 319..321
 FT /tag- c
 FT /note- "Encodes Ala"
 XX
 PN W0200172842-A2.
 PD 04-OCT-2001.
 PD 28-MAR-2001; 2001WO-US09996.
 PF 28-MAR-2001; 2000US-192978P.
 XX
 PR 28-MAR-2001; 2000US-192978P.

Db 2041 CTGTTTGATGATCAGCTCAGCGGCCAGCTGCTTATCTGTCTAATCTGGCTGTGTG 2100
 QY 1591 CGGCTTACGTCACACCCCGTACTGACCCCGATGACCCCAAGATCAATTTGCTCTGT 1650
 Db 2101 TGGACCCACATGCTGCTAGGGAATACAGGCTTCCCGCATCTGTATGCTTATGTC 2160
 QY 1651 AACCCCACTACCGGACAGCGCTGTCTTCAACACAGCTGAGCTCTGCTCTCACTG 1710
 Db 2161 ACAGAGACCACTCCCTGGGCTTCACTGAGCTTCTCTCAATAGGCTCTCTCAATC 2220
 QY 1711 GTGGTTACGCTTGGCTTACATGAGGCAAGAGCTGCCACCACTACAGAGGCGAAG 1770
 Db 2221 AGTGGCTTGGCTTGGATCTGAGCTGAGGCTTCCAGAGAACTCAACAGAGGCGAAA 2280
 QY 1771 TTGATCACCCTCAGCATACCTTCTATTTCACTTCACTCCGCTCTGACCTTCATG 1830
 Db 2281 TGTGCTACCTTACGCTCTCTCTTCACTTGTGCTGATGCTTCTTCAACAGGCGC 2340
 QY 1831 TCTGCTTACAGCGGGGCTGTGTCACATGCTGACCTTGTGCTACCTGTCTCAACCTC 1890
 Db 2341 AGCGTCTACGAGCGAATACCTGCTGCGGCAACATGATGAGCTGGCTGAGCAGCTG 2400
 QY 1891 CTGGGCAATCAGCTGGCTTACTTGGCGCCCAAGTGTACATGATCTTCTTACCCGAG 1950
 Db 2401 AGCAGCGCTTGGTGGTATTTCTGCTTAACTGCTACGATGATCTTCTGCGCGCCAGC 2460
 QY 1951 CGCAACAGCGCCGCTTACTTCAACAGATGATCCAGGCTACACCATGAGAGGAGCT 2008
 Db 2461 CTCACAGCAGACAGACACTTTCAGGCTTCATTCAGAGCTACAGAGGCGCTGCGGCT 2518

RESULT 7 AAD19505 standard; DNA; 2951 BP.

XX AAD19505;
 XX AAD19505;
 DT 18-DEC-2001 (first entry)
 DE Human novel G-protein coupled receptor DNA with 5' and 3' regions.
 XX
 KW Human; G-protein coupled receptor; GPCR; gene therapy; drug screening;
 KW antiseizure therapy; signal transduction; behavioural disorder; obesity;
 KW heartbeats rate; inflammation; immune disorder; diabetes; cancer;
 KW coronary disease; ds.
 XX
 OS Homo sapiens.
 XX
 FN WO200172842-A2.
 PD 04-OCT-2001.
 XX
 PE 28-MAR-2001; 2001WO-US09996.
 XX
 PR 28-MAR-2000; 2000US-192978P.
 XX
 PA (LEXI-) LEXICON GENETICS INC.
 PI Walke DW, Scoville J, Donoho G, Turner CA, Friedrich G, Abuh A;
 PI Zambrowicz B, Sands AT;
 XX
 DR WPI; 2001-616474/71.
 PT Novel isolated polynucleotides encoding human G protein coupled
 PT receptor (NGPCR), useful for detecting mutant NGPCRs or inappropriately
 PT expressed NGPCRs for diagnosis of disease, and as probes or primers -
 XX
 PS Disclosure; Page 79-80; 80pp; English.
 CC
 CC The present sequence is human novel G-protein coupled receptor (NGPCR)
 CC DNA with 5' and 3' regions. NGPCRs are transmembrane proteins that span
 CC the cellular membrane and are involved in signal transduction after
 CC ligand binding. The NGPCR polynucleotide sequences, are useful

CC in diagnosis and treatment of a disease involving NGPCR, for detecting
 CC mutant NGPCRs or inappropriately expressed NGPCRs for the diagnosis of
 CC disease, for screening drugs effective in treatment of symptomatic or
 CC phenotypic manifestations of perturbing the normal function of NGPCR in
 CC the body or abnormalities in the signal transduction pathway mediated by
 CC NGPCR. NGPCR DNA is also useful in gene therapy, drug screening, clinical
 CC trial monitoring and/or the treatment of physiological (heartbeats rate)
 CC or behavioural disorders. NGPCR is useful for identifying compounds
 CC useful in the therapeutic treatment of obesity, inflammation, immune
 CC disorders, diabetes, heart and coronary disease, metabolic disorders and
 CC cancer. Labelled NGPCR nucleotide probes can be used to screen a human
 CC genomic library which is helpful for identifying polymorphisms,
 CC determining the genomic structure of a given locus/allele and designing
 CC diagnostic tests.

SO Sequence 2951 BP; 559 A; 888 C; 831 G; 669 T; 4 other;

Query Match 19.28; Score 386.4; DB 22; Length 2951;

Best Local Similarity 51.18; Pred. No. 1.6e-71;

Matches 1042; Conservative 1; Mismatches 947; Indels 48; Gaps 4;

QY 1 ATCACTTACAGCGGATGAGTGTGAGAGCAAGTGGCTTCCGCTTGTGTG 60
 Db 730 ATTAAGTATGCGGCGACGACGACGCTTCAAGCTGAGCGGCAATCTCTTCTCTG 789
 QY 61 CGTACACACCCAGCGCGACACACGTCGAGGCAATGTCAGCTGATCTGACCTTC 120
 Db 790 CGCACATCCCAATGACATGACATGACAGTGTGAGACATGATGCTGCTGCAAGATGTC 849
 QY 121 CGCTGGAATGATATGCTGTGTGTGAGCAGCAGACATGATGCGCGGCAATGCGCAG 180
 Db 850 GGGTGGACCTGGAATCTCTGTGTGTGAGCAGATGATGAGGAGCTGAGGAGGCTGAG 909
 QY 181 CTGCTTGGCGAGCGGTGCGCGCGGCGGCAATGATGATGCTTCCAGAGACCTGCGCC 240
 Db 910 GCACTGGAGAACGACGCGACCTGCTGAGGAGATCTGATCTTTCACAGACATATATGCC 969
 QY 241 AACTGACGCCCAACACAGACATGATGATGAGCGCGCGCTGATGATGATGTC 300
 Db 970 -----TTCCTGCGCCAGGTGGGCGATGAGAGATGATGATGCTCATG 1011
 QY 301 GACAACTGACAGCAGAGACAGCGCGGCTGTGTGTGCTGTCTGCGCGGACCTGACCTG 360
 Db 1012 CGCCACCTGCGCGCGCGCGCGCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTG 1071
 QY 361 TACCACTTCTCAATGATGATGCTGTGCGGCAATGATGATGATGATGATGATGATGATG 420
 Db 1072 AGGCTGTTTTCGAGTCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1131
 QY 421 GAGTCTGGGCGCATGACCGCGGTCTGTGACACATGATGATGATGATGATGATGATGATG 480
 Db 1132 GAGGCTGGGCGCTCTCCAGCAGATCATCTGGGTGCGCGGATGATGATGATGATGATG 1191
 QY 481 TTCTTGGGCAATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540
 Db 1192 GTGCTGGGCGGTGCGATCAGAGAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1251
 QY 541 GGCCACAGGCTGGGCG 594
 Db 1252 TATGCGCGGCGAGACAGAGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1311
 QY 595 AACCAAGAGTGGCAACTGCTGATGAGCGCACTTGTCTTCAACACCATTTCTAGGCTC 654
 Db 1312 AATCAGCTCTCAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1371
 QY 655 TCTGGGAGGCTG 714
 Db 1372 TTCTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1431
 QY 715 CACAAGCTCTGCGCTGTGACAAAGACACTGCAACAAGAGGCTGTCTACCTGTGAGAG 774
 Db 1432 CACCAAGCTCTGCGCTGTGAGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG 1491

QY	775	CTGCTGATGAGGAGATCTGGAAGGATGACATCTCACCTCTGGAACACCAAAATCTCTTGAC	834
Db	1492	CTTTTGGAGAGATTCACCAAGGTGCACTTTCCTCTTACACAAAGAACACTGTGGCTTAAAT	1551
QY	835	CCGCAAGGGAGACCTGGCTCTGCACTTGTGGAGATTGTCCATGTGGCAATGGAGACGAGCCAG	894
Db	1552	GACAAACAGAGATCCCTCACTACCTATTAACATAATTGCCCTGGAGACTGGAAATGACCCAG	1611
QY	895	AATCCCTTCCAGAGCGTGCCTCTTAC-----TACCCCTGCAGCGACAGCTGAAGAC	948
Db	1612	TGGACCTTTCACGGTCCCTGGTCTCTCCACATGGGTCCAGTTTACACTAAACATAATATAG	1671
QY	949	ATCAAGACATCTCTGCACACCCCTCAACAAACAGATCCCATATCTCCATGTCTTCCAAAGG	1008
Db	1672	ACCAAAATCCAGTGGCACGGAAAGGACCAAGGTGCCATMACTGTGTGTTCAGCGAC	1731
QY	1009	TGCAGTCAGGGCAAAAGAAAGACCTGTGGGCATCACTGTGTGGCTGTGAGTGCATC	1068
Db	1732	TGCTTTGAAGGGCACCAAGCGAGTGTAGGGTTTCATCACTGCTGCTTTGAGTGTG	1791
QY	1069	GACTGCTTCCCGGACCTTCTCAACGACACTGA-----TGCCG	1110
Db	1792	CCCTGTGGGCGCTGGAGACCTTCCCAACAGAGTGAACCTTACAGATGCGACGCTGTGG	1851
QY	1111	AATTAAGAGTGTCTTACACGATGTAGACCTCTGCTTAAACGGCAGCTGTCTCTG	1170
Db	1852	AAAGAAAGTGTGGACCTGTAGGGAGCCAGACACTGTGTTCCCGGACCTGTGTGTTTTTG	1911
QY	1171	GAATGGATGTAGGACCCACACATCGCTGTGGCCCTCTGGCCGCTGTGGCTCTCTAGC	1230
Db	1912	GCTTTGGGTAGACACACTTGTGGGTGTGTGTGGACAGCTAACACGCTGTGTGCTGTG	1971
QY	1231	ACCTGGCACTTCGTGTGATATTCTGGAAGCACTTCAGACACCAATATGTGTGCTGTG	1290
Db	1972	CTGCTTGGGACTGTGCGCTGTGCTTGGCTGTGGACCTTGAACCCCTGTGGTGAAGTCACA	2031
QY	1291	GGGGGGCCCAATGTCTCCGATGTGACACACTGTGTGTGTGGCATACATGTGTGCCG	1350
Db	2032	GGGGGGCCGCTGTCTTCTTAATGTGGGCTCCCTGTGGACAGATGTGGACGCTCTAT	2091
QY	1351	GTGTACCTGTGGGGCGCCCAAGGTCTCACCTGTGCTGTGGCGCAGGCGCTTTCCTCT	1410
Db	2092	GGCTTCTTTGGGAAACCCACAAAGCCTGTGTGTCTTACGCGCAAGCCCTCTTTCCTT	2151
QY	1411	TGCTTCACAATTTGCATCTCTGTATCGCGTGTCTTCTTTCAGATGTGTGCGCTTC	1470
Db	2152	GCTTTCACACTTCTCGTCTCGCTGTGACAGTTCGGTCAATTCACACTATCATCATCTTC	2211
QY	1471	AAGATGGCCAGCGCTTCCACGCGCTTACAGCTACTGGGTCCGTACCAAGGGCGCTTAC	1530
Db	2212	AAGTTTTCACCAAGATACCTTACATTCCTACACGCGCTGGGTCCAAACACAGTGTCTGGC	2271
QY	1531	GTCTCTATGAGCATTTATACAGGTACTCAAAATATGTGTATGTGGCATGTGTGCA	1590
Db	2272	CTGTTTGTGATGATCACTGTAGCGGCCACGTGTCTTATGTCTTAACTTGGCTGTGTG	2331
QY	1591	CGGCGCTAGTCCACCCCGCTACTGTGACCCCGATGACCCCAAGATCACAATTGTCTCTGT	1650
Db	2332	TGGAACCCACGTGCTGTAGGAATACAGAGGCTTCCCATCTGTGTATGTGTGATGTGAGTGC	2391
QY	1651	AACCCCCACTTACGCAACACCTGTGTTCACACACAGCTGTGACCTGTGTCTCAAGT	1710
Db	2392	ACAGAGACCAACTCCCTGGGCTTCAATACGTGGCTTCTCTCAAAAGGCTCCTCTCAATC	2451
QY	1711	GTGGTTTCAAGCTTGGCTGTACATGTGGGCAAAAGCTCCCAACACTACAAAGAGGCCAAG	1770
Db	2452	AGTGCCTTTGCTGTGCACTACTGTGGGTAAAGATTTGCCAGAGACTTACAAAGAGGCCAAA	2511
QY	1771	TTTCAACCCCTACAGATGACTTCTTATTTACCTCATCTCGTCTCCCTGTGCACCTTCAATG	1830
Db	2512	TGTGTACCTTCAAGCCGCTGCTTCAACTCTGTGTCTGTGATGTGCTTGTTCACACAGGCC	2571
QY	1831	TTCTGCTTACAGCGGGGTGTGTGTCACCAATGTGACCTTGTGTGCTACTGTGTCAACTTC	1890

Db	2572	AGCTCTACGACGGCGCAAGTACCTGCGGCGCCAAATATATGCTGAGCGCTG	2631
Qy	1891	CTGGCCATCAGCCCTGCGGCTACTTCGCGCCCAAGTGTACATGATCCTCTTCTACCCGAG	1950
Db	2632	AGCAGCGGCTTCGCTGGGTAATTTCTGCGCTTAAGTGTACGTATCCTCGCGCCGAC	2691
Qy	1951	CGCAACACGCGCCCTCTTCAACAGCATGATCCAGGCTACACCATGAGAGGACT	2008
Db	2692	CTCAACAGCAGACGAGCACTTCACAGGCTCCATTCAGAGACTACAGACGAGCGCTG	2749
RESULT 8			
XX	ABL55953		
XX	ABL55953 standard; cDNA; 2759 BP.		
XX	ABL55953;		
XX	17-JUN-2002 (first entry)		
XX	Human G-protein coupled receptor encoding cDNA SEQ ID NO 13.		
DE	Human: GPCR; G-protein coupled receptor; receptor; anti-HIV; antitumour;		
KW	antihistaminergic; antiallergic; antianemic; antiasthmatic; antiviral;		
KW	immunosuppressive; dermatological; nephrotoxic; antipruritic; antihypertensive;		
KW	cytostatic; neuroprotective; osteoporotic; antiparasitic; antineoplastic;		
KW	antiarthritic; thyromimetic; antifungal; ophthalmological; antibacterial;		
KW	fungicide; antiparasitic; protozoacide; antihelminthic; antidiabetic;		
KW	antiartherosclerotic; hepatotropic; anticonvulsant; anorectic; metabolic;		
KW	antimetic; antidiarrhoeic; neuroleptic; cerebroprotective; nootropic;		
KW	antiparkinsonian; depilatory; tranquilizer; hypotensive; vasotropic;		
KW	cardiac; antianal; vulvular; proliferative disorder; cancer;		
KW	neurological disorder; Alzheimer's disease; Huntington's disease;		
KW	Parkinson's disease; multiple sclerosis; meningitis; prion;		
KW	cardiovascular disorder; acquired immunodeficiency syndrome; AIDS;		
KW	Crohn's disease; diabetes mellitus; rheumatoid arthritis; gene; ss.		
XX	Homo sapiens.		
OS			
FH	Key	Location/Qualifiers	
FT	CDS	53..2578	
FT		/*tag= a	
FT		/product= "G-protein coupled receptor"	
XX	WO200198323-A2.		
XX	27-DEC-2001.		
XX	15-JUN-2001; 2001WO-US19354.		
XX	16-JUN-2000; 2000US-212483P.		
PR	23-JUN-2000; 2000US-213950P.		
PR	26-JUN-2000; 2000US-214062P.		
PR	07-JUL-2000; 2000US-216585P.		
PR	14-JUL-2000; 2000US-218936P.		
PR	19-JUL-2000; 2000US-219154P.		
PA	(INCY-) INCYTE GENOMICS INC.		
XX	Lal P, Graul R, Hafalia AJA, Walla NK, Thornton M, Nguyen DB;		
PI	Lu Y, Gandhi A, Patterson C, Kallilick DA, Baughn MR, Ramkumar J;		
PI	Tribooley CM, Lee EA, Ding L, Burford N, Yao MG, Yang J;		
PI	Griffin JA;		
XX	WPI: 2002-1139780/18.		
DR	P-PSDB; ABB77319.		
XX	Novel G-protein coupled receptor protein and polynucleotides useful for		
PI	diagnosing, treating or preventing disorders of cell proliferation e.g.		
XX	cancer, neurological and genetic disorder e.g. thalassemia		
PS	Claim 50; Page 117-118; 121pp; English.		
XX			

CC The invention relates to a G-protein coupled receptor protein (GCREC)
 CC polypeptide (AB77317-AB77326). The GCREC is useful for screening an
 CC agonist/antagonist of GCREC, a compound that specifically binds to GCREC
 CC or that modulates the activity of GCREC. GCREC is also useful as an
 CC immunogen for preparing antibodies which are useful for diagnosing a
 CC condition of disease associated with expression of GCREC in a subject,
 CC for detecting and purifying GCREC from a sample. The GCREC encoding
 CC polynucleotide (AB75951-AB75960) is useful for screening for a compound
 CC effective in altering expression of GCREC. GCREC is also useful for
 CC assessing toxicity of a test compound useful for treating a disease or
 CC condition associated with decreased expression or overexpression of
 CC functional GCREC. Examples of disorders include cell proliferative
 CC disorder such as arteriosclerosis, atherosclerosis, hepatitis, mixed
 CC connective tissue disease (MCTD), psoriasis and cancer including
 CC adenocarcinoma, leukemia; a neurological disorder such as epilepsy,
 CC stroke, Alzheimer's disease, Huntington's disease, dementia, Parkinson's
 CC disease, retinitis pigmentosa, multiple sclerosis, bacterial and viral
 CC meningitis, abscess, subdural empyema; prion disease including kuru,
 CC Creutzfeldt-Jakob disease; fatal familial insomnia, neurofibromatosis,
 CC tuberous sclerosis, cerebral palsy, polymyositis; inherited, metabolic,
 CC endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis;
 CC mental disorders including mood, anxiety, and schizophrenic disorders;
 CC seasonal affective disorder (SAD); akathisia, amnesia, catatonia,
 CC diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses,
 CC Tourette's disorder; cardiovascular disorders such as hypertension,
 CC vasculitis, vascular tumours, congestive heart failure, ischaemic heart
 CC disease, myocardial infarction, calcific aortic valve stenosis,
 CC infective endocarditis, endocarditis of systemic lupus erythematosus,
 CC cardiac transplantation; gastrointestinal disorder such as dysphagia,
 CC gastritis, anorexia, nausea, emesis, abdominal angina, infections of
 CC the intestinal tract, peptic ulcer, hepatitis, cirrhosis, diarrhoea,
 CC acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, Reye's
 CC syndrome, liver infarction; an autoimmune/inflammatory disorder such as
 CC AIDS, Addison's disease, adult respiratory distress syndrome,
 CC allergies, amyloidosis, anaemia, asthma, atherosclerosis, autoimmune
 CC thyroiditis, bronchitis, Crohn's disease, diabetes mellitus,
 CC Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis,
 CC multiple sclerosis, myasthenia gravis, myocardial or pericardial
 CC inflammation, osteoporosis, psoriasis, rheumatoid arthritis,
 CC scleroderma, Sjogren's syndrome, systemic lupus erythematosus, uveitis,
 CC viral, bacterial, fungal, parasitic, protozoal, and helminthic
 CC infections, and trauma; a metabolic disorders such as diabetes,
 CC obesity, and osteoporosis. The GCREC encoding polynucleotide is useful
 CC for gene therapy and for creating knock in humanised animals (pigs) or
 CC transgenic animals (mice or rats) to model human disease.

SQ Sequence 2759 BP; 523 A; 828 C; 782 G; 626 T; 0 other;

Query Match 19.28; Score 385.2; DB 24; Length 2759;

Best Local Similarity 51.18; Pred. No. 2.9e-71;

Matches 1042; Conservative 0; Mismatches 948; Indels 48; Gaps 4;

QY 1 ATCACTACAGCGCATGAGTGGAGAGAGTGGCGCTTCCGCTTTGGTG 60
 DB 551 ATTAGCTATGGGCGGAGAGAGAGCTTCAAGCGGCGGATATCTCTTTCTG 610
 QY 61 CGTACACACCCAGCGCGACACACGATCGAGGCGATGTGACAGTGTCTCATCTT 120
 DB 611 CGCACATATCCCATGACAGTACAGGTGAGACATGCGTCTCTCTCAAGATTC 670
 QY 121 CGGTGAACTGATCTGTGTGTGAGAGAGACACTATGGCGCGACATGGCCAG 180
 DB 671 GGGTGGACTGTGATCTGTGTGTGAGAGAGACTATGGCGGAGTGGGTGAG 730
 QY 181 CTGTGGCGGAGCGGCGCGCGGAGACATGTGATCGGCTTCCAGAGAACCTGGCC 240
 DB 731 GCACTGGGAACAGCGGCGGAGGAGATGTGATCTTTCATGAGCATATGCCC 790
 QY 241 ACACTGAGCCCAACAGATACGTCAAGAGAGCGCGCGCTGTGTGACATTTGTG 300
 DB 791 -----TTCTCTGCCAGGTGGGCGATGAGAGATGCAATGCTCTATG 832
 QY 301 GACAAGCTGACAGACAGCGCGGCGTGTGCTGTCTCTGCGCCGACCTGACCTG 360

DB 833 CGCCACCTGGGCGGCGGCGGCGGAGACCGTGTGTTTCTTTTTCACCGGCAATGGCC 892
 QY 361 TACACCTTCTTCAATGAGGTCTGCGCCAGACACTTACGCGGCGCTGTGATGGCTCC 420
 DB 893 AGGGTGTCTTTCGAGTCCGTGGTGTGACCAACATGCTGTGCAAGGTGGGTGCTCA 952
 QY 421 GAGTCTGGGCGCATGACCGCGGTCTCTACACACTCAGAGAGTGGGCGCACTTGGGCA 480
 DB 953 GAACCTGGGCGCTCTCCAGGACATCATCTGAGGTGGCGGAGATCCAGCCATGGGATG 1012
 QY 481 TTCTGGGCATACCATTCACAGAGGTGCCATCCCGGCTTCACTAGTATCCGAGTGG 540
 DB 1013 GTCTGGGCGGCGCATATCCAGAGAGGCTCTCTCGCTGAGGGCTTGAAGAGCC 1072
 QY 541 GGCCACAGGCTGGGCGGCGGCGGCGGCGGAGAGAGGAGGAGGAGTAT-----ACCTG 594
 DB 1073 TATGCGGGGAGACAGAGAGAGGCGGCTTGAAGGCTTGCACAGGCGCTCTGTGAGCAG 1132
 QY 595 AACGAGAGTGGCAGCACTGCTTGAACGCACTTGTCTTCAACACCATTTCTAGGCTC 654
 DB 1133 AATAGCTCTGACAGATGCAAGATGCAAGCTTTATGAGCAGACAGATGCCAAGCTCAAGCC 1192
 QY 655 TCTGGGAGGCTGTCTCTACAGGCTTACTTGGCTTATGCTGTGGCCATGCTCTG 714
 DB 1193 TTCTCATGATGTTCTGCTTACAGCATATACGAGTACGCGGCTGTATGCGGCGCATGCGCTC 1252
 QY 715 CACAGCCTCTCGGCTGTGACAAAGACACCTGACACCAAGGAGGTGTACCCCTGGGCG 774
 DB 1253 CACAGCTCTCGGCTGTGCTGTGAGCTGTGAGCTGTTCACAGGCGGCGAGTCTTACCCCTGGGCG 1312
 QY 775 CTGCTTGAAGAGATCTGGAAGTCACTACTCTCTCTGAGACCAACCAATCTTCTGAC 834
 DB 1313 CTTTGGAGAGATTCACAAAGGCTCATTTCTTACAAAGAGACTGTGGGCTTAAAT 1372
 QY 835 CCGCAAGGAGAGCTGCTCTGCACTTGGAGATTTCCAGTGGCAATGGAGCGGACG 894
 DB 1373 GACAACAGAGATCCCTCAGTAGTATTAATATTCCTGGGCTGGAATGGACCCAG 1432
 QY 895 AATCCCTTCCAGAGCTGCTCTCTAC-----TACCCCTGCGACGACACTGAGAAC 948
 DB 1433 TGAACCTTCAAGCTCTGCTGCTCTCTCTCCACATGCTCTCAAGTCAAGCTTAAATATGAG 1492
 QY 949 ATCAAGACATCTGCAACACCGCTCAACACAGATCCCTATGTCTATGTCTCCAGAG 1008
 DB 1493 ACCAAATTCAGTGGAGGAGAAAGACACAGGCTGCTAATGTGTGTCTTCCAGCGAC 1552
 QY 1009 TGCCAGTCAAGGCAAAAGAGAGAGCTGTGGGCACTCCAGCTGTGCTCTGAGTGCATC 1068
 DB 1553 TGTCTTGAAGGCGACAGCGAGTGTATAGGGTTCATCATCTGCTTGTGAGTGTGTG 1612
 QY 1069 GACTGCTTCCGCGACCTTCTCTCAACACACTGAA-----TGCCCG 1110
 DB 1613 CCTGTGGGCGTGGGACTTCTCTCAACAGAGTACTCTTACAGATGCGACGCTGTGGG 1672
 QY 1111 AATAGAGAGTGTCTCTACAGAGTGAAGACTCTCTGCTTCAAGCGGAGCTGTCTCTG 1170
 DB 1673 AAAGAGAGTGGGAGCTGAGGAGAAAGCAGACCTGCTTCCGCGACACTGTGTGTTTGG 1732
 QY 1171 GAATGCAATGAGGACACCACTGCTGTGTGGCTCTGCTGGCGCGCTGTGGCTTCTAGC 1230
 DB 1733 GCTTGGGTGAGCACTCTTGGGTCTGTGGCACTTAACAGAGCTGTGCTGTCTGTG 1792
 QY 1231 ACCCTGGCATCTGAGATATTTGTGAGACACTTTCAGACACCACTATGTTGCTGGCT 1290
 DB 1793 CTGCTTGGAGCTGTGCTGCTGTGTGCTGTGCACTTAACACACCTGTGTGAGTGA 1852
 QY 1291 GGGGCGCCATGCTCTCTGATGCTGACACTGCTGTGTGTGAGCATATGTTGTCCG 1350
 DB 1853 GGGGCGCGCTGTCTTCTTATGCTGTGGGCTCCCTGCAAGAGATGAGGAGCTCTAT 1912
 QY 1351 GTGTAGTGGGCGGCGGAGGCTCTCACTGCTGTGCGGCGAGGCGCTTCTTCCCTC 1410


```

Db 1913 GGCCTCTTTGGGGAACCAACAGGCGCTGCTGTTGCTAGCAGGAGCCCTCTTGGCCTT 1972
QY 1411 TGTTCACAAATTGATCTCTGATATGCGCCGTGTTTTCAGATGCTGCGCCTTC 1470
Db 1973 GGTTCACAACTTCTCTGCTGCGTACAGTGTGCTTCACTATCATCATCTTTC 2032
QY 1471 AAGATGGCCAGCGGCTCCCAAGCGGCTACAGCTACGCTGGGTCGCTACAGGGGCCCTAC 1530
Db 2033 AAGTTTCCACCAAGTACATCATTTCTACACAGCCTGCTGCAAAACACCGGTGCTGCG 2092
QY 1531 GTCCTATGAGCATTTATCAGGCTACTCAAAATGGTCAATGTGTAATGCGATGCTGCA 1590
Db 2093 CTGTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2152
QY 1591 CGGCTCAGTCCACCCCGTACGACCCGATGACCCGATGACCCGATGACCCGATGACCC 1650
Db 2153 TGGACCCCACTGCGCTGAGGGAATACGAGCGCTTCCCAATGCTGATGCTGATGATG 2212
QY 1651 AACCCCACTACCGGACACAGCCTGCTGTTCAACACAGCCTGAGACCTGCTGCTCACTG 1710
Db 2213 ACAGAGACCAACTCCCTGCGCTTCTACTGCGCTTCTCTCAATGCGCTTCTCTCAATC 2272
QY 1711 GTGGTTTCACTTCCCTACATGAGGCAAGAGCTGCCACCACTACATGAGGCAAG 1770
Db 2273 AGTGGCTTGGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG 2332
QY 1771 TTCACTACACCTCAGATACCTTCTATTTCACTACCTGCTGCTGCTGCTGCTGCTG 1830
Db 2333 TGTGTCACCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2392
QY 1831 TCTGCTCAGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1890
Db 2393 AGCGCTCAGAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2452
QY 1891 CTGGCCATCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1950
Db 2453 AGCAGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2512
QY 1951 CGCAACAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2008
Db 2513 CTCAACAGCAGAGCAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2570

RESULT 9
AA250745
ID AA250745 standard; cDNA: 2333 BP.
XX
AC AA250745;
XX
DT 31-MAY-2000 (first entry)
XX
DE Human sensory transduction G-protein coupled receptor-B3 cDNA.
XX
KW Human; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;
KW sensory cell; taste receptor cell; screen; taste modulator;
KW pharmaceutical; food industry; taste topographic map; tongue; ss.
XX
OS Homo sapiens.
XX
FH Key 1..2333 Location/Qualifiers
FT CDS /*tag= a
FT /product= "GPCR-B3"
FT /transl_except= (pos:865..866, aa:His)
FT
PN MO200006592-A1.
XX
PD 10-FEB-2000.
XX
PE 27-JUL-1999; 99MO-US17099.
XX
PR 28-JUL-1998; 98US-009465.

```

```

PA (REGC ) UNIV CALIFORNIA.
PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Zuker CS, Adler JE, Lindemeier J, Ryba N, Hoon M;
XX WPI: 2000-205451/18.
XX P-PSDB: MAY45023.
DR
DR New isolated sensory transduction G-protein coupled receptor, useful
PT for developing products for use in studying and modulating the taste
PT transduction pathway -
XX
XX Claim 5; Pages 79-80; 83pp; English.
XX
CC The present sequence is a cDNA encoding a taste cell specific G-protein
CC coupled receptor, GPCR-B3 which is involved in sensory transduction.
CC This sequence was isolated from the human testis library.
CC GPCR-B3 is specifically expressed in foliate and fungiform cells, with
CC lower expression in circumvallate taste receptor cells of the tongue.
CC The present sequence is used to screen compounds that modulate sensory
CC signalling in taste cells, especially taste modulators useful in
CC pharmaceutical and food industries to customise taste. The sequence
CC can also be used as probe for identifying taste cells and
CC subsets of taste receptor cells such as foliate, fungiform and
CC circumvallate. Such probes are also useful to generate taste
CC topographic maps that elucidate the relationship between the taste
CC cells of the tongue and sensory neurons leading to taste centres
CC in the brain.
XX
XX Sequence 2333 BP; 445 A; 705 C; 646 G; 537 T; 0 other;
XX
XX
XX Query Match 18.7%; Score 376.6; DB 21; Length 2333;
XX Best Local Similarity 50.9%; Pred. No. 1.7e-69;
XX Matches 1036; Conservative 0; Mismatches 954; Indels 47; Gaps 4;
XX
QY 1 ATCACTCAGAGGCGCATGAGTGGGAGAGAGGCGCTTCCGCGCTTGGTG 60
Db 307 ATTACCTATGGGGCCAGAGGAGAGAGCTCAGCGTGAAGCGGCAATCCCTTTCTG 366
QY 61 CGTACACACCCAGGCGGACACGACGATGAGGCGATGGTGACATGCTGCATTC 120
Db 367 CGCACATCCCAATGATGACAGTACGAGTGGAGACATGCTGCTGCGAAGATTC 426
QY 121 CGCTGAGACTGATATGCTGCTGAGAGAGAGGACCTATGCGCGGAGCAATGGCCAG 180
Db 427 GGGTGGACCTGATCTCTGCTGTTGGAGAGTACGATGAGGAGGAGCTAGGGGCTGAG 486
QY 181 CTGCTTGGAGGCGGCGGCGGCGGAGATCTGATGCGCTTCCAGAGAGAGCGTGGCC 240
Db 487 GCACGTGAGAGACAGCCCGCTGTCAGGGGCACTGCTATGCTTAAAGACATCATGGCC 546
QY 241 ACACTGACGCCCAACAGACATGACGTGAGAGAGCGCCAGCGCTGCTGATGCTG 300
Db 547 -----TTCTCTGCCAGGCGGCGGAGTGAAGAGATCAGAGGAGCTCATG 588
QY 301 GACAAGCTGACAGACAGAGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 589 CGCACCTGGCCCAAGGCGGCGGCGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 648
QY 361 TACCACTTCTCAATGAGTGTGCTGCGCAGAACTTCAAGGGCGGCGTGTGATCCCTCC 420
Db 649 AGGCTGTTTTCGAGTCCGTGTGTGACAACTGACTGCAAGATGTGGTCCCTCA 708
QY 421 GAGTCTGGGCGCATGACCCGCTCTGCAACAACCTCAGAGAGTGGGCGCACTTGGGCA 480
Db 709 GAAGCTGGGCGCTCTCCAGGCACATCACTGGGCTGCGCGGAGATCAGCATTTGGATG 768
QY 481 TTCTGGGCGCATCAGATCAGAGGCTGCCAATCCGGGCTTCAAGATTCGGAGTGG 540
Db 769 GTGCTGGGCGTGGCGCATCAGAGAGGCGTGTCCCTGCGTGAAGGCGTTGGAAGAGCC 828
QY 541 GGCCACAGCGTGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 595

```

Db 829 TATCCCGGGGAGACAAAGAGGCCCCCTAGGCTTGACACAGGGGCTCTGTGACAGACAGA 888
 QY 596 ACCAGAGAGTGGAGAACTGCTGAAAGGCACTTGTCTTCAACAGCAATCTCAAGGCTCT 655
 Db 889 ATCAGCTCTGAGAAATGCCAGCTTTCATGAGACACAGATGCCAGCTCAAAAGCT 948
 QY 656 CTGGGAGACGTGTCTACAGCGCTTACTCTGCGGTCTATGCTGTGGCCATGCGCTGC 715
 Db 949 TCTCCATGAGTCTGCTGCAACAGCATACCGGCTGTATGCGGTGGCCATGCGCTGC 1008
 QY 716 ACAGCTCTGCTGGGTGACAAAAGCAACCTGACCAAGAGGGTGTCTACCCCTGGGACG 775
 Db 1009 ACAGCTCTGCTGGGTGACAAAAGCAACCTGACCAAGAGGGTGTCTACCCCTGGGACG 1068
 QY 776 TGTCTGAGAGATCTGAAAGGTCAACTCTCTCTGAGACCAACAAATCTCTGACG 835
 Db 1069 TTTTGAGACATCTCAAGAGTGTGATTTCTCTTACACAAAGACACGTGTGGCGTTTATG 1128
 QY 836 CGCAAGGGAGCGTGTCTGCACTTGGAGATTTTCAGTGGCAATGGGACCGAGCCAGAGA 895
 Db 1129 ACAACAGAGATCCCTCAGTACATTAACATTAATGCTCTGGAGCTGGAATGACCCAGT 1188
 QY 896 ATCCCTTCCAGAGGCTGCTCTCTAC-----TACCCCTGAGGAGAGCTGAAGAACA 949
 Db 1189 GGACCTTACAGGCTCTCTGCTCTCTCTCAGATGCTCTCAGTTCAGTAAACATTAATGAGA 1248
 QY 950 TCAAGACATCTCTGACACCTCTCAACACAGATCCCTATGCTCATGTGTCCAAGAGT 1009
 Db 1249 CCAAAATCCAGTGGCGGAAAGAAACACAGTGCCCTAATGTGTGTCCAGCGACT 1308
 QY 1010 GCCAGTACGGGCAAAAGAAAGCGCTGTGGGATCCAGCTGTCTGCTCTGATGATG 1069
 Db 1309 GTCCTGAAGGGGACACAGAGAGTGTATGCGGTTTCCATCACTGCTGTGAGTGTGTC 1368
 QY 1070 ACTGCTTCCCGGACCTTCTCTCAACCACTGAA-----TGCCCGA 1111
 Db 1369 CCTGTGGGCTGGAGACTTCTCTCAACAGAGCGAGCTTACAGATGCCAGCTGTGGA 1428
 QY 1112 ATAAAGAGTGTCTTACAGAGTGAAGACCTCTGCTTCAAGCGGAGCTGTCTCTG 1171
 Db 1429 CAGAAGAGTGGGACCTGAGGAGAGCCAGACCTGCTTCCGCGACGTGTGTGTTTGG 1488
 QY 1172 AATGAGATGAGGACCAACCACTGCTGTGGGCTGTGGGCTGTGGGCTGTCTCTGACA 1231
 Db 1489 CTTTGGGCTGAGACACCTTGGGTGTGCTGGAGAGCTTACACACCTGTGCTGTGTC 1548
 QY 1232 CCTTGGGCTGCTGTGATTTCTGAGAGCACTTCCAGACACCAATAGTGTGCTGTG 1291
 Db 1549 TGTCTGGGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1608
 QY 1292 GGGGCCCCATGTGCTGTGATGTGACACTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1351
 Db 1609 GGGGCCCCATGTGCTGTGATGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1668
 QY 1352 TGTACGTGGGCCCCCAAGGTCTCACCTGCTGTGCGGCGAGCCCTCTTCCCTCT 1411
 Db 1669 GCTTCTTGGGGAACCAAGAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1728
 QY 1412 GCTTCACAATTTGATGCTGCTGTATGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1471
 Db 1729 GTTTCACATTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1788
 QY 1472 AGATGGCCAGCGCTTCCACAGCGCTTACAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1531
 Db 1789 AGTTTCCACCAAGGTACCTTATACCAAGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1848
 QY 1532 TCTCTATGAGCTTATACAGGTACTCAAAATGTCTATGTTGTAATTGGATGTGCTGAC 1591
 Db 1849 TGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1908
 QY 1592 GGGCTCATGCTCCACCCGCTGTGATGACCCGATGACCCCAAGATACACATTTGTCTCTGTA 1651
 Db 1909 GGACCCACATGCTGTGTAGGGAATACAGCGCTTCCCGCATCTGTGTATGCTGTGATGCA 1968

QY 1652 ACCCACTACCGCAACAGCGTGTCTTCAACACCAGCTGTGCTGTGCTGTGCTGTGCTGTG 1711
 Db 1969 CAGAGACCAATCCCTGTGGCTTACTGAGCTTCTCTCAAAATGGCTCTCTCTCAATCA 2028
 QY 1712 TGGGTTTACGCTTGTCTACATGTGGCAAGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1771
 Db 2029 GTGCTTGTGCTGTGAGTGTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2088
 QY 1772 TCATCACCCTGAGTGTGCTGTGATTTTCACTATGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1831
 Db 2089 GTGTACCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 2148
 QY 1832 CTGCTACAGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1891
 Db 2149 GCGTCTACAGAGGCAAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 2208
 QY 1892 TGGGCTATGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1951
 Db 2209 GCAGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 2268
 QY 1952 GCACAGCGCGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 2008
 Db 2269 TCAACAGCAGACACTTCCAGGCTTCCATTCAGAGCTACAGAGCGCGCTGCGGCT 2325

RESULT 10
 ID AAD17516 standard; cDNA; 2526 BP.
 AC AAD17516;
 AC 10-DEC-2001 (first entry)
 DT 10-DEC-2001 (first entry)
 DE Human taste receptor, hT1R1 cDNA coding sequence.
 KW Human; taste-cell-specific G protein-coupled receptor; hT1R1; drug;
 KW genetic modulation; pharmaceutical; taste sensation; food industry;
 KW chemosensory transduction; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 1..2526
 FT /tag- a
 FT /product- "Human taste receptor, hT1R1 protein"
 FT /transl_except- (pos:820..823, aa:Phe)
 FT /note- "this codon has an apparent 1 nucleotide insertion
 which alters the reading frame"

WO200166563-A2.
 13-SEP-2001.
 07-MAR-2001; 2001WO-US07265.
 07-MAR-2000; 2000US-0187546.
 07-APR-2000; 2000US-0195536.
 06-JUN-2000; 2000US-0209840.
 23-JUN-2000; 2000US-0214213.
 17-AUG-2000; 2000US-0226448.
 03-JAN-2001; 2001US-0259227.
 (SENO-) SENOMYX INC.
 Adler JE, Zozulya S, O'Connell SM, Li X, Staszewski L;
 WPI; 2001-582267/65.
 P-PDB; AAE10372.
 New mammalian taste-cell-specific G protein-coupled receptor
 polypeptides for identifying compounds that modulate taste signaling
 are useful in food, to modulate the sweet taste of foods or drugs

xx
ps Claim 1; Page 82-83; 119pp; English.
xx
CC The invention relates to mammalian taste-cell-specific G protein-coupled
CC receptors, T1R and their corresponding cDNA molecules. Taste receptors,
CC T1R are useful for screening compounds which are used to activate or
CC modulate chemosensory transduction, such as taste sensation. The
CC identification and isolation of novel taste receptors and taste
CC signalling molecules allow for new methods of chemical and genetic
CC modulation of taste transduction pathways. The taste modulating
CC compounds are useful in pharmaceuticals and food industries to improve
CC the taste of a variety of consumer products, or to block undesirable
CC tastes, e.g., in certain pharmaceuticals. T1R's are also useful in
CC biochemical assay for identifying tastant (T1R) ligands having binding
CC specificity for T1R involved in taste signalling. The present cDNA
CC sequence is human taste-cell-specific G protein-coupled receptor, hT1R1
CC coding sequence.
xx
SO Sequence 2526 BP; 470 A; 770 C; 697 G; 589 T; 0 other;

Query Match 18.3%; Score 368.8; DB 22; Length 2526;
Best Local Similarity 50.9%; Pred. No. 7.7e-68;
Matches 1038; Conservative 0; Mismatches 952; Indels 48; Gaps 5;

OY 1 ATCACTTACAGCCGATGAGTGCAGACAAAGTCCGCTCCGCTTTCGTCG 60
DB ATTACTATGCGGCGAGGAGAGGCTCAGCGTGAAGCGGAGATCCCTCTTCG 558
OY 61 CGTACCAACCCAGCGCGGAGCAGCAGTGCAGGCCATGCTGAGTGTGCACTTC 120
DB CCGACCATCCCAATGACAAATGACAGGAGGAGACCATGCTGCTGTCGAGAAATTC 618
OY 121 CGCTGGAATGATGATGCTGAGGAGCAGCAGTATGAGCGGCGGAGCAATGCGCAG 180
DB GGGTGAAGCTGATCTCTGCTGAGAGAGTGAAGTATGAGGAGGAGGAGGAGGAG 678
OY 619 GGGTGAAGCTGATCTCTGCTGAGAGAGTGAAGTATGAGGAGGAGGAGGAGGAG 678
OY 181 CTGCTTGGGAGAGCGGCGGCGGAGCAGTATGAGTGCAGTGCAGGAGAGAGCGTCC 240
DB GCAGTGGAGAAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 738
OY 241 ACAGTGAAGCCCAACAGAACTGACGAGAGAGAGGAGGAGGAGGAGGAGGAGGAG 300
DB TCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
OY 301 GACAAGCTGACAGAGCAGAGCAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAG 359
DB CGCCACAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 840
OY 360 GTACCACTTCTCAATGAGGAGTCTGCGGAGAACTTACAGGAGGAGGAGGAGGAGG 419
DB CAGGGGTGTTTTCGAGTCCGTGCTGACCAACTGACTGAGGAGGAGGAGGAGGAGG 900
OY 420 CGAGTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 479
DB AAGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
OY 901 AAGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 960
OY 480 CTTCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 539
DB GGTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1020
OY 540 GGGCCCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 594
DB TATGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1080
OY 595 AACCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 654
DB AATCAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1140
OY 655 TCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 714
DB TTTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1200
OY 715 CACAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 774

DB 1201 CACAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1260
OY 775 CTGCTTGAAGAGATGAGAGGATCAACTTCACTCTCTGAGGAGGAGGAGGAGGAGGAG 834
DB CTTTGGAGGAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1320
OY 835 CCGCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 894
DB 1321 GACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1380
OY 895 AATCCCTTCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 948
DB 1381 TGGAGCTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1440
OY 949 ATCAAGAGATCTCTGAGCAGGCTCAACAGCAGATCCATGATGATGATGATGATGATG 1008
DB 1441 ACCAAATTCAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500
OY 1009 TGGCAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1068
DB 1501 TGTCTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1560
OY 1069 GAGTGCCTTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1110
DB 1561 CCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1620
OY 1111 AATTAAGAGTGTCTTACAGAGTGAAGTCTCTGCTTCAAGCGGAGGAGGAGGAGGAGG 1170
DB 1621 AAGAAGAGTGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680
OY 1171 GAATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1230
DB 1681 GCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
OY 1231 ACCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1290
DB 1741 CTGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
OY 1291 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1350
DB 1801 GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1860
OY 1351 GTGTAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1410
DB 1861 GCTTCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1920
OY 1411 TGCCTTCAAAATTTGATCTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1470
DB 1921 GGTTCACCATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
OY 1471 AAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1530
DB 1981 AAGTTTTCACCAAGGATCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTT 2040
OY 1531 GTCTATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1590
DB 2041 CTGTTTGTATGATCACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2100
OY 1591 CGGCTCACTCCACCCCGTACTGAGCCGATGAGCCCAAGATCACAATTTCTCTCTGT 1650
DB 2101 TGGAGCCCACTCTCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
OY 1651 AACCCCACTTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1710
DB 2161 ACAGAGACCAATCTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220
OY 1711 GTGGGTTTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1770
DB 2221 AGTGCCTTTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2280
OY 1771 TTTATCACCCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1830


```

OY 1010 GCCAGTCAGGGCAAAAGAGAGCGTGTGGCATTCCAGCTGTGCTGTGAGTGCATCG 1069
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1515 GTTCGAGAGGGCACACAGAGTTGTATGGTGTCCACACACTGTCTGTGAGTGCATGC 1574
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1070 ACTGCTTCCCGGACCTTCTCAGCCACATGAA-----TGCCCGA 1111
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1575 CCGTGAAGTGGGACATTTCTCAGCAGAGTAGCTTCACACTGTCCAGCCCTTGTGGAA 1634
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1112 ATACAGAGTGTCTTACAGAGTAGAGACTCTGTCTTCAAGGGGACAGCTGTCTCTGG 1171
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1635 CACAAGATGTGGGCCCTGAGGGAGGCTCAGCTCTCTCAGCAGCCGTGAGATTTCTTGG 1694
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1172 AATGGCATGAGGACCAACCATGCTGTGGCCCTGTGCGGCCCTGTGGGCTTCTCAGCA 1231
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1695 GGTGGCATGAGACCATCTCTTGTGTCTATTAGACGCTAACAGCTATTTCTCTGTCTGC 1754
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1232 CCCGTGCATCTGTGTATATCTGTGAGGACATTCAGACACCCATATGTCGTGGCTG 1291
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1755 TGATTTGGAGTGTGTGGGCTGTGCGGTGTCTTACAGCCGTGTGTGAGGTGACGCTG 1814
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1292 GGGGCCCCATGTCTTCTGTATGCTGACACTGTCTGTGTGGCATATGTTGTCGGG 1351
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1815 GGGGTAGGCTGTCTCTCATCTGTGGTTCCTTGTGTAGCTGGAGTTGACGCTCTACA 1874
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1352 TGTACGTGGGGGCGCCCAAGGTCTCCACGCTGCGCTGCGCGCCCTTCTTCCCTCT 1411
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1875 GCTTCTTGGGAAAGCCACAGGTGCGCGGTGTCTGTCTGACGCTCTTCTTCTTCTG 1934
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1412 GCTTCACAATTTGCATCTCTGTATGCGCGCTGTCTTCTTCAAGTGTGTGCGCTTCA 1471
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1935 GGTTTGCAATTTCTCTCTCTGTGTGACATTCGCTCTTCAACGTGTATCATCTTCA 1994
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1472 AGATGGCCAGCCGTTTCCACAGCGCTACAGCTACTGTGGTCCGTACCAAGGGCCCTAG 1531
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 1995 AGTTTCTACCAAGATCCACATCTACACACTGTGGGCCCAAAACATGAGTGCAGAA 2054
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1532 TCTCTATGCGATTATACAGGTACTCAAAATGTCATTTGTATTTGGATGTGGTGCAC 1591
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2055 TATTTGCTATTTGCTACAGTCTCCAGGTCATTTGTTCTCTCTCTCAGTGTGCAATGT 2114
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1592 GGCCTGATGCCACCCCGTACTGACCCGATGACCCGATGACCCCAAGATCATAATGTCTCTGTA 1651
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2115 GGAACCCACAGGGCCACAGGAGTACAGCGCTTCCGCCATTCGTGTATTTGAGTGC 2174
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1652 ACCCCAACTACCGCAAGAGCTGTCTGTCAACACAGCTGTGAGCTGTCTCTCACTGG 1711
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2175 CAGAGGTCAACTGTGTGGGCTTCTGTGTGCTTTCGCACAAACATCTCTCTCCATCA 2234
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1712 TGGGTTTACAGCTTCGCTTACATGGGCAAGAGCTGGCCACCACTACATCAAGAGGCAAT 1771
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2235 GCACCTTTGTCTGAGTACTCTGTGTAGAGTAATGCGGAGAACTATACAGAAAGCCAAAT 2294
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1772 TCATCAACCTCAGCATGACCTTCTATTTTACACATCCGTCCTCTGTGACCTTCAATGT 1831
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2295 GTGTCACTTCTCAGCTGTCTCTCACTCTGTATTCGTGTGATTCGCTTCTTCTTCAATGTGCA 2354
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1832 CTGCTTACAGAGGGGTGTGTGTGTACCATCTGTGACCTCTGTGTGTGTCTCACTCTC 1891
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2355 GCATTTTACAGAGGGGACCTACTACCCGCGGTCAATGTGTGGAAGGGCTGTGCACTCTGA 2414
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1892 TGGCCATTCAGACCTTGGCTACTTGGGCCCCAAGTCTACATACATCTTCTTACCCGGAGC 1951
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2415 GTGGCGGCTTACAGCGGTATTTCTCTTAAATGTATAGTATCTCTGTCTGTCCAGAAC 2474
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 1952 GCAACAGCGCCGCTACTTCAACAGCATGATCCAGGGCTTACACATGAGGAGGAGAC 2007
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2475 TCAACAACAGAACTTTCAGGCTCCATCCAGAGACTACAGAGAGCGTGGGCGC 2530
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 12
 AAS46935
 ID AAS46935 standard; cDNA; 697 BP.

```

XX AC AAS46935;
XX DT 18-DEC-2001 (first entry)
XX DE Human G protein-coupled receptor (GPCR) cDNA #117.
XX XX
KW Human; G protein-coupled receptor; GPCR; mental disorder; schizophrenia;
KW neurological disorder; metabolic disorder; cancer; rheumatoid arthritis;
KW thyroid disorder; neurodegenerative disorder; cardiovascular disorder;
KW renal failure; autoimmune disorder; hyperproliferative disorder; HIV; ss;
KW human immunodeficiency virus; viral infection; neuroprotective;
KW anorectic; gene therapy.
XX XX
OS Homo sapiens.
XX XX
PN WO200168858-A2.
XX XX
PD 20-SEP-2001.
XX XX
PF 16-MAR-2001; 2001MO-US08456.
XX XX
PR 16-MAR-2000; 2000US-187783P.
PR 16-MAR-2000; 2000US-189907P.
PR 16-MAR-2000; 2000US-189917P.
PR 16-MAR-2000; 2000US-189918P.
PR 16-MAR-2000; 2000US-189960P.
PR 29-MAR-2000; 2000US-192155P.
PR 29-MAR-2000; 2000US-192234P.
PR 29-MAR-2000; 2000US-192830P.
PR 29-MAR-2000; 2000US-192916P.
PR 29-MAR-2000; 2000US-192933P.
PR 29-MAR-2000; 2000US-192935P.
PR 29-MAR-2000; 2000US-192945P.
XX XX
PA (PHEA ) PHARMACIA & UPJOHN CO.
XX XX
PI Vogel1 G;
XX XX
DR WPI; 2001-607458/69.
XX XX
DR P-PSDB; AAU29496.
XX XX
PT Nucleic acid encoding G protein-coupled receptors, useful for the
PT prevention, diagnosis and treatment of mental disorders -
XX XX
PS Claim 4; Page 97; 274pp; English.
XX XX
CC Sequences AAS46819-AAS46946 represent cDNA molecules encoding human G
CC protein-coupled receptor (GPCR) polypeptides. The protein and DNA
CC sequences of the invention can be used to identify compounds which bind
CC to GPCR polypeptides and in screening for compounds that modulate GPCR
CC activity. By screening a human subject for the presence of mutations in
CC GPCR DNA, a GPCR-related disorder or a genetic predisposition can be
CC diagnosed. The sequences can also be used for treatment and prevention of
CC mental disorders such as schizophrenia, neurological disorders such as
CC manic depression, metabolic disorders such as obesity, cancer, rheumatoid
CC arthritis, thyroid disorders such as myxoedema, neurodegenerative
CC disorders such as Parkinson's disease, cardiovascular disorders such as
CC atherosclerosis, renal failure, autoimmune disorders, hyperproliferative
CC disorders such as psoriasis and viral infections such as those caused by
CC HIV.
XX XX
SQ Sequence 697 BP; 130 A; 177 C; 211 G; 179 T; 0 other;
XX XX

```

Query Match 12.8%; Score 258; DB 22; Length 697;
 Best Local Similarity 100.0%; Pred. No. 9.6e-45;
 Matches 258; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 1105 TGCCCGAATAACGAGTGTCTTACAGAGTGAAGCTCTCTGCTTCAAGGGCAGCTGTGTC 1164
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 440 TGCCCGAATAACGAGTGTCTTACAGAGTGAAGCTCTCTGCTTCAAGGGCAGCTGTGTC 499
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```

Db 1859 GCGCTCTCTGCTCCCTGGCCAGCCAGCCCTGCCAGTGGCTGGCCAGAGCCCTGT 1918
QY 1403 TTCCCTCTGCTTACATTTGCAATTCCTCTATGCGCGCTCTTTTCCAGATGCTCT 1462
Db 1919 CCCACCTCCCGCTACGCGGCTGAGACACATCTTCTGACAGCGCGCCGAGATCTTCG 1978
QY 1463 GCGCTCTCAAGATGGCCAGCCGCTTCCACAGCGCTTACAGTACTGGGCTGCTACAGCAG 1522
Db 1979 TGGAGTAGAAGATGCTCTGAGCTGGGCGAGACCGGCTGAGTGGCTGCTGGGGGCGCT 2038
QY 1523 GGGCTTACGCTCTATGAGCAATTTATCAAGGTACTCAAAATGTCATTGTGTAATTGCA 1582
Db 2039 GGGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2098
QY 1583 TGGCTGAGACGCGCTTACATCCACCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1642
Db 2099 TGGCTGAGCTTCCCGCGGAGGTGCTGAGACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2158
QY 1643 TCTCTGTAACCCCACTACCGCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1702
Db 2159 TGCACCTGCGGACACGCTTCTGAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2218
QY 1703 TCTGAGTGGTGGTTCATGCTTCCCTACATGGCCAAAGAGCTGCCACCACTACAAAG 1762
Db 2219 TGGCTTCTCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2278
QY 1763 AGGCAAGTTCATCAGCTTACAGTACAGCTTCTATTTACCTGCTGCTGCTGCTGCTGCTG 1822
Db 2279 GTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2338
QY 1823 CCTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1882
Db 2339 CCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2398
QY 1883 TCAACCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1942
Db 2399 TCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2458
QY 1943 ACCGAGAGCGACAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2458
Db 2459 AGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

```

RESULT 15
 AAL38462 standard; cDNA, 3458 BP.
 AAL38462;
 AAL38462;

15-AUG-2002 (first entry)
 cDNA encoding a human G-protein coupled receptor (GPCR).

G-protein coupled receptor; GPCR; gene chip; human; immune response;
 chromosomal position; transgenic animal; gene therapy; gene; ss.

Homo sapiens.

Key location/Qualifiers
 CDS 30..2588
 /tag- a
 /product- "Human G-protein coupled receptor"

W0200230981-A1.

18-APR-2002.

13-MAR-2001; 2001WO-US07832.

10-OCT-2000; 2000US-0684393.

(PEKE) PE CORP NY.

XX
 PI Wei M, Zhong W, Ketchum KA, Difrancesco V, Beasley EM;
 XX WPI, 2002-444173/47.
 DR P-PDB; AAO21501.
 XX
 PT Novel G protein coupled receptor, useful for raising antibodies, to
 PT elicit immune response, and as a reagent in assays designed to
 PT quantitatively determine levels of protein in biological samples -
 XX
 PS Claim 1; Fig 1; 82pp; English.

The invention relates to an isolated G-protein coupled receptor (GPCR)
 CC polypeptide, comprising an 852 residue amino acid sequence, given in the
 CC specification, an allelic variant or ortholog of the protein, or a
 CC fragment comprising at least 10 contiguous amino acids of the protein.
 CC GPCR is useful for identifying a modulator of GPCR and an agent that
 CC binds to GPCR. GPCR and a gene chip comprising GPCR are useful as models
 CC for the development of human therapeutic agents. GPCR is useful for
 CC raising antibodies, to elicit immune response, as a reagent in assays
 CC designed to quantitatively determine levels of protein in biological
 CC samples, and as markers for tissues in which the corresponding protein
 CC is preferentially expressed. A gene chip containing GPCR is also useful
 CC as a probe for determining the chromosomal positions of nucleic acid
 CC molecules by means of in situ hybridisation, in making vectors containing
 CC the gene regulatory regions of a gene chip containing GPCR, for designing
 CC ribozymes, in making vectors that express GPCR, and for constructing host
 CC cells and transgenic animals expressing nucleic acid molecules and
 CC peptides. A host cell containing GPCR is useful for conducting cell-based
 CC assays involving GPCR protein or its fragments, and for identifying
 CC protein mutants in which these function is affected. The polynucleotide
 CC encoding GPCR can be used to treat disorders by gene therapy. This
 CC polynucleotide sequence represents cDNA encoding the human G-protein
 CC coupled receptor of the invention.

XX
 SQ Sequence 3458 BP; 590 A; 1179 C; 1092 G; 597 T; 0 other:

Query Match 10.9%; Score 218.6; DB 24; Length 3458;

Best Local Similarity 46.8%; Pred. No. 2.6e-36;

Matches 941; Conservative 0; Mismatches 1014; Indels 54; Gaps 6;

```

QY 2 TCACCTACAGCGCCATAGAGATAGCTGCGAGCAAGGTGCTTCCCGCTTGTCTGC 61
Db 523 TCACCTACAGCGCTAGAGAGCTGAGCGCGCCGAGAGACTTCCCTCTCTTCTCC 582
QY 62 GTACACACACCGAGCGCGACACAGCTGAGCGCTGAGCTGAGCTGAGCTGAGCTGAG 121
Db 583 GCACCGTCCCGAGCGAGCTGAGCTGAGCGCGCGCGAGCTCTGAGAGATTGC 642
QY 122 GCTGGAAGTGGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 181
Db 643 GCTGGAAGTGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 702
QY 182 TGTCTGCGAGCGGCTGCGCGCGAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 241
Db 703 TGTCTGCGCGCTGCGCGCGAGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 762
QY 242 CACTGACGCCCAACAGACATGAGCTGAGAGCGCGCGAGCGCTGCTGCTGCTGCTGCTG 301
Db 763 TGCCCGTGGCGATGATCGCGCGCTGCGGAGAGTGCAGAGCTCTC----- 810
QY 302 ACAAGCTGACAGACAGACAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361
Db 811 ACCAGGTGAACAGACAGAGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 870
QY 362 ACCACTTCTTCAATGAGGTGCTGCGGAGAACTTACAGCGCGCGCTGCTGCTGCTGCTGCG 421
Db 871 AGCGCTCTTCAATGAGCATGAGCATGAGCATGAGCATGAGCATGAGCATGAGCATGAGCATG 930
QY 422 AGTCTGCGCGCATGAGCGCGCTGCTGCGAACAACCTGAGAGCTGAGGACCTTGGGCACT 481
Db 931 AGGCTGCTGAGCTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

```

QY 482 TCCTGGGCAATCCATCCAGAGCGTCCCATCCGGGCTTACAGTATCCCGAGTGG 541
 Db 991 TGTTCGGCTTCCCTCCAGAGGGGGTCCAGTGTACAGATTCCTCCGATGATGAACGC 1050
 QY 542 GC-----CCACAGGCTGGGCGCCACCCCTCAGAGGACGACGACCTATACC 591
 Db 1051 ACCTGGCCCTGGGCGACCGACCCGCTTCGTCTGCTCCCTGGGCGAGAGGACAGGGTTC 1110
 QY 592 TGCACACGAGAGGACGACAACTGCTGAAAGCCACTTGTCTTCAACACATCTTCAGG 651
 Db 1111 TGGAGAGAGGACGTGTGGGCGACGCTGCCCGCATGTATGATCATCACGTGACAAAG 1170
 QY 652 CTCTTGGGAGCGTGTCTGTACAG-----CGTACTCTGGGCTATGCTGTGG 703
 Db 1171 TGGAGGAGGCGGCTAAATACACACGATCTCTGTCTACAGAGCTGTGTATAGGTGG 1230
 QY 704 CCCTATCCCTGCACAGCCTCTCTGCGCTGTGACAAAAGACCTGACACCAAGAGGG---TGG 760
 Db 1231 CCCAGGCCCTGCGACAACTCTTCACTGTGCAACGCTCAGGCTTGCCTCCGCGACGACCCG 1290
 QY 761 TCTACCCCTGGGAGCTGTGTAGAGATCTGGAAGGTCAACTTCTCTCTGACACAC 820
 Db 1291 TGAAGCCCTGGGAGCTCTGGAGAACATGTACAACTGACCTTCCACGTGGGCGGCTGC 1350
 QY 821 AAATCTTCTGACCCGCAAGGGGAGGTGCTGTGACCTTGGAGATTGTCCAGTGGCAAT 880
 Db 1351 CGCTGGGTTTCGACACACACGAGGAAAGCTGTGACATGTGATACGACTGAAAGTGTGGGTGT 1410
 QY 881 GGGACCGGAGCCAGAAATCCCTTCCAGAGCGTGCCTCTCACTACCCCTGACGACGACAG 940
 Db 1411 GGCAGGGCTCAGTGCACAGGCTCCACGAGGTGGGAGGTTCAACGGGACGACTAG---GA 1467
 QY 941 TGAAGAACATCAAGACATCTCTGCAACACCTCAACAAACAGATCCCTATGTCTATGTGT 1000
 Db 1468 CAGAGGCGCTGAAGATCCGCTGGCACACGCTGTGACAAACGAAAGCCGTTCCGGTGGCT 1527
 QY 1001 CCAAGAGTGCAGTGCAGTGCAGGCAAAAGAAAGAGCTGTGGGCAATCCACGTCGTGCTTCG 1060
 Db 1528 CGGCGGCAAGTGCAGGAGGCGGAGGTGCGCGGGGTCAAGGGGTTCCACTCTGTGTGCTACG 1587
 QY 1061 AGTGCATGCAGTGCCTTCCCGGACCTTCTCAACCACTGAA----- 1104
 Db 1588 ACTGTGTGACTGCGAGGGGCGGAGCTACCGGCAAAACCGAGACATGCGCTGCACCT 1647
 QY 1105 --TGCCGAATACAGAGTGGTCTTACACAGTGAAGACCTCTGCTCAAGCGGACACTGG 1162
 Db 1648 TTTGTGGCCAGATGTAGTGTCTCCCGGAGCGAAAGCACAGCTGTCTCCGCGCAGGTCTC 1707
 QY 1163 TCTTCTGGAATGGCATGAGGACCCACCATGCTGTGGCCCTGTGAGCCGCTTGGGCT 1222
 Db 1708 GGTTCCTGGATGGGGCGAGCGGCTGTGTGCTGTGCTGTGCTGTGAGAGCTTGGCGG 1767
 QY 1223 TCTTCAGACACCTTGGCCATCTGTGTATTTCTGAGGACATTCAGACACCATATGTTTC 1282
 Db 1768 TGGGCTTGTGTGCTGT 1827
 QY 1283 GCTGGGCTGGGGGCCCATGTGTCTGTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1342
 Db 1828 AGGCTCTGGGGGGGGCCCTGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1887
 QY 1343 TGGTCCCGGTGTAGTGGGGCGCCCAAGGTCTCACTGTGCTTCCGCGCAGGGCCCTCT 1402
 Db 1888 GGTTCCTGTCTGTCTGT 1947
 QY 1403 TTTCCCTGT 1462
 Db 1948 CCACACTCCGCTCAGGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2007
 QY 1463 GCGCTTCAAGATGGCGAGCGGCTTCCAGCGGCTTACAGACTAGGGGTCCGCTACAGG 1522
 Db 2008 TGGAGTCAAACTGCTCTGTAGTGTGGGCGAGACGCGGTGTGTGTGTGTGTGTGTGTGT 2067
 QY 1523 GGCCCTACGTCTGTATGGATTTATACGGTACTCAAAATGGTCAATGTGTATTTGGCA 1582

Db 2068 GGGCTGTGGCTGT 2127
 QY 1583 TGTGTGACGGCTTCACTGATCCACCCCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1642
 Db 2128 TGTGTGCTTGT 2187
 QY 1643 TGT 1702
 Db 2188 TGT 2247
 QY 1703 TGT 1762
 Db 2248 TGT 2307
 QY 1763 AGGCAAGTTTATCATCCCTGACATGACCTTCTATTTCACCTTCACTGCTGTGTGTGA 1822
 Db 2308 GTGCCGCTGGGCTTCACTTGTGCTATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2367
 QY 1823 CCTTCAATGT 1882
 Db 2368 CCTCTGTGCGCAATGT 2427
 QY 1883 TCAACTCTGTGGCATCAGCTGTGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1942
 Db 2428 TGT 2487
 QY 1943 ACCCGAGGCGCAACAGCGCCGCTACTTC 1971
 Db 2488 AGCGAGGCTCAACACCCCGGAGTCTTC 2516

Search completed: May 23, 2003, 13:07:43
 Job time : 476.125 secs

THIS PAGE BLANK (USPTO)

Query Match Similarity 7.7%; Score 154.2; Db 17; Length 1014;
 Best Local Similarity 50.0%; Pred. No. 7e-22;
 Matches 435; Conservative 1; Mismatches 430; Indels 4; Gaps 2.

QY	1105	TECCCCGAATACGAGTGTGCTACCAAGATGAGACCTCTGCTCAAGCGGACGTGTC	1164
Db	875	TTGTGAAAGACACCATGTGGTCCACAGCTGGAGACACTCTCTKCTCCACAGGGTGGAG	816
QY	1165	TTCCGGAATGAGCATGAGGACCCACCAATGCGTGTGGCCCTGC-TGGACGGCCGTGGCTT	1223
Db	815	TGGGTGCCGCTTACAGACACCCCGCGGCTGTGTATCTCTGCGCGGGGCTGCTTGTGCT	756
QY	1224	CCTCAGCACCCCTGGCCATCTGGTATATTTTGGAGGACCTTCAGACACCCATAGTTG	1283
Db	755	GGCCCTTACACCGACGACATGTGGTCTCTCTGCGCTCACTCAACACACCGCGTGGTCG	696
QY	1284	CTCGGCTGGGGGCCCCCATGTCGTCTCTGATGCTGACACGCTGTGTGGCATATGCT	1343
Db	695	ATCTCCGCGGGGACCCATGTGCTCTCTCAATCTTAAGCTGCTCCAGCCTTGTAGATCAG	636
QY	1344	GGTCCCGGTATACGTGGGGCCGCCAAGGTCTCCACCTGCTCTGCGCGGACGCCCTT	1403
Db	635	CGCTCTCTTTTACTTGTGAGAGACGACATGAAACCTTTTGTATCTTAGAGTTATGCGCTT	576
QY	1404	TGCCCTGTGCTTACAAATTTGCAATCTCTGTATATGCGCCGTGGCTTCTTCCAGATGCTG	1463
Db	575	TCTTTTATTTCTTACGCTGCTGTCTGTGCGCTGTTTCCGCTCCCTCTCTTCAATGTCTAT	516
QY	1464	GGCCTTCAAGATGGGACAGCGCTTCCACGGCGCTCCACACTGCTCTGCGCGGACCGG	1523
Db	515	CATTTTAAATATAGCCGCGCAAGTTCGCCGAGTCCACAGCTGTGTGATATAGTACCAGG	456
QY	1534	GGCCTACGCTCTCTATGCAATTTATCAAGGTACTCAAAATGTCTATGTGTATTTGGCAT	1583
Db	455	CCAGTGGCTGTGCATCTCCATGACCTTTGTCTCTTACGAGCATTTGATTTGATGGGCTT	396
QY	1584	GCTGGCAGGGCTCAGTCCACCCCGTATGACCCCGATGACCCCAAGATCAATTTGT	1643
Db	395	CTCTGTCCAAACCCACCTTGTCCCTACANGRCRTCTCTCTTACCCAGACAAATCATCT	336
QY	1644	CTCTGTAAACCCCACTACCGACAGCGCTGTGTTCAACACAGCCTGGACCTGTGCT	1703
Db	335	GGGCTGTGACGCTTA--ATCTCAATATGATCATGAGCTATTTTCTTCTCTTATTTGTT	279
QY	1704	CTCAGTGGGGGTTTCAAGTTCGCTTCACTATGGGACAGAGGCTGGCCACCATACAGA	1763
Db	278	GTGCATCTCTCTCTTACTTCTCTTACATGGGGAAAGACTCTCCGAGAACTACACGA	219
QY	1764	GGCAGATTTGATCACCCCTCAGCATGACCTTCTATTTTACCTCATCCGTCCTCCCTGAC	1823
Db	218	GGCCAAAGCATTAACCTTCTGCTGTCTGTGATCTTCACTCGATCATATTTGGCCAC	159
QY	1824	CTTCATGTCTGCTTACAGGGGGTGTGTGTACCATCTGTGGACCTCTTGTGTACATGTGT	1883
Db	158	GGCTTTATGTCTTACACAGGCAATATACATCCACACCTTCAAGCGCTGGCGTCTTC	99
QY	1884	CAACCTCTTGGGCATCAGCTTGGGCTACTTTCGCCCCAAGTCTCATATGCTCTCTTA	1943
Db	98	CAGGCGCTTACTCTTCTCTCTGTGTGTATTTCTCCCAATGTGTATATATATCTTTCA	39
QY	1944	CCCGGAGGCAACAGCCCGGCTACTTCAA	1973
Db	38	GGCTTCACAAGAACCCCAAGAACTACTTCCA	9

RESULT 4
 CDS03010/c
 LOCUS CDS03010 979 bp DNA linear GSS 15-MAY-2000
 DEFINITION Tetradon nigriviridis genome survey sequence PUC-Orl end of clone
 18412 of library G from Tetradon nigriviridis, genomic survey
 sequence.
 ACCESSION AL222381

VERSION	KEYWORDS	AL222381.1	GI:7881200
SOURCE	ORGANISM	Genome survey sequence.	
		Tetraodon nigroviridis.	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.	
REFERENCE	TITLE	1 (bases 1 to 979)	
	AUTHORS	Roeest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C., Bernot,A., Flzames,C., Wincker,P., Brottier,P., Quetier,F., Saurin,W. and Weissbach,J.	
REFERENCE	TITLE	2 (bases 1 to 979)	
	AUTHORS	Roeest-Crollius,H., Jallion,O., Dasilva,C., Flzames,C., Fisher,C., Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and Weissbach,J.	
		Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis	
JOURNAL	TITLE	Unpublished	
REFERENCE	TITLE	3 (bases 1 to 979)	
	AUTHORS	Genoscope.	
		Direct Submission	
		Submitted (12-APR-2000)	
COMMENT		This sequence is a single read and was generated as part of a large scale clone-and-sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/tetraodon .	
FEATURES	Source	Location/Qualifiers	
		1..979	
		/organism="Tetraodon nigroviridis"	
		/db_xref="taxon:99883"	
		/clone="184120"	
		/clone_1bp="G"	
		/note="Genoscope sequence ID : COAG184DF10SP1-end : PUC-Or1"	
BASE COUNT	227 a	230 c	299 g 198 t 25 others
ORIGIN			
Query Match	5.7%	Score 114;	DB 17; Length 979;
Best Local Similarity	49.4%	Pred. NO. 1.8e-13;	
Matches 339;	Conservative 12;	Mismatches 332;	Indels 3; Gaps 3;
OY 1143	CTGCTTCAAGGGGAGCGCTGGTCTTCCCTCGAATGGATGAGGACCCACCAATCGCTGTGGC	1202	
Db 669	CTKTCGTGAGGAGACTCGGCTGCMATCTGGATATGGACCATCTTATGTGGTGGCTCTCT	630	
OY 1203	CTGTGTGCGCCGCCCTTCTCTGAGCACCCTGGCCATCTGTGATATTTCTGGAGCA	1262	
Db 629	TTTCTTCTGTGTGCTGTCTTCTCATGACCTCCAGCTCGGAGATGATCTCTGCTCAA	570	
OY 1263	CTTCCAGACACCCTAATGTTCCTCGGCTGGGGGGCCCAATGTCTTCTGATGTGACACT	1322	
Db 569	CATTAACACGGCTGTGGGCAAGTCGGCGGAGGCGACCTGCGCCCTTAAAGCTGGGCGC	510	
OY 1323	GCTGTGTGTCATATGATGTGTCCCGTATAGCTGGGGGGCCCAAGGATCTCCACTG	1382	
Db 509	TCTAAGACAGACAGCAATGACTCTGTGTGACACTTTGGCCAGCCGCTCCCGCTGGCCAG	450	
OY 1383	CCCTGTGCGCCAGGCGCTCTTCCCTCTGTCTGCTTCAACAATTTGATCTCTG-TATGCGCG	1441	
Db 449	CAAGCTCAAGACGGCTCAATTCAGTTTCAAGTTCACTGTGTCTGCTGCAATTCGCTG	390	
OY 1442	TGGCTTTTTCATATGCTCTGGCGCTTCAAGATGGCCAGCCGCTTCCACAGCGCTTACA	1501	
Db 389	TGGCGTCCGACAGGATCAATTCATTTTAAATTCGCGCTGCGGTGC-CGCGACGCTTACA	331	
OY 1502	GCTACTGTGCTCCGCTTCAAGGAGGCGCTTCAATGCTATGATGATTTATCAAGGTAGTACAAA	1561	
Db 330	ACAGTGTGGCCAAAACCATGTGGCGCTTCAAGGAGGCGCTTCAATGCTATGATGATTTATGATGATGCTGCGGCAATCC	271	

QY	1562	TGTCATATGTTGTAATTTGGCAAGGCGGACGCGCTCAAGTCCACCGGCTGACACCGG	1621
Db	270	TGTGGCTGTCTGTCTGTGGCCGTAAGCCGTC - GGCCCCCCCCAGCCCTCCARRATCTTATT	212
QY	1622	ATGACCCCAAGATCAATCAATGTTCTCTGTAAACCCCACTACCGCAACAGCCCTGCTGTCA	1681
Db	211	TCMACAGCAACAGATATCGMGCTGRAGMGACACAAACACCTCTCGCCGGTTCCTTGCG	152
QY	1682	ACACCAAGCCCTGACCTGCTGCTCTCAAGTGTGGTTCACGTTTGCCCTACATGGGCAAG	1741
Db	151	AGCTCTGTAAAGCTAGCTGTGCTGAGCGCCGTTTGCTGTTTGAAGMACATGGGCAAR	92
QY	1742	AGTCCGCCACCAACATGACAGAGGCGCAATGATCACCCTCAGCATGACCTTATATTCA	1801
Db	91	ACGTGCGGCGCACTACACAGAGCGCAAGTGCGTACCTGATGCTGATGGTGAACATRA	32
QY	1802	CCATCCGCTCTCCCTGTCGACCTTC	1827
Db	31	TCTCCTGGATTAGCTTCTTCAACCGTC	6
RESULT 5			
CNS02BUU			
LOCUS			
DEFINITION	CNS02BUU	973 bp	DNA
KEYWORDS	Tetraodon nigroviridis genome survey sequence PDC-Or1 end of clone		
SOURCE	254011 of library G from Tetraodon nigroviridis, genomic survey		
ORGANISM	sequence.		
ACCESSION	AL190335.1	GI:7828439	
VERSION	GSI: genome survey sequence.		
KEYWORDS	Tetraodon nigroviridis		
SOURCE	Tetraodon nigroviridis		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;		
AUTHORS	Tetraodontidae; Tetraodon.		
REFERENCE	1 (bases 1 to 973)		
AUTHORS	Roeest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,		
REFERENCE	Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,		
AUTHORS	Saurin,W. and Weissenbach,J.		
REFERENCE	Human gene number estimate provided by genome wide analysis using		
AUTHORS	Tetraodon nigroviridis DNA sequence		
REFERENCE	Unpublished		
AUTHORS	2 (bases 1 to 973)		
REFERENCE	Roeest-Crollius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C.,		
AUTHORS	Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and		
REFERENCE	Weissenbach,J.		
AUTHORS	Characterization and repeat analysis of the compact genome of the		
REFERENCE	Freshwater pufferfish Tetraodon nigroviridis		
AUTHORS	Unpublished		
REFERENCE	Genoscope.		
AUTHORS	3 (bases 1 to 973)		
REFERENCE	Submitted (12-APR-2000)		
AUTHORS	This sequence is a single read and was generated as part of a large		
REFERENCE	scale clone-end sequencing project of the Tetraodon nigroviridis		
AUTHORS	genome. For more information, please take a look at		
REFERENCE	http://www.genoscope.cns.fr/Tetraodon.		
FEATURES			
SOURCE	Location/Qualifiers		
	1..973		
	/organism="Tetraodon nigroviridis"		
	/db_xref="taxon:99883"		
	/clone="254011"		
	/clone_1bp="G"		
	/note="Genoscope sequence ID : COAG254AH06SP1-end :		
	PDC-Or1"		
BASE COUNT	184 a	306 c	247 g
ORIGIN			227 t
			9 others
Query Match	4.9%	Score 97.8:	DB 17:
Best Local Similarity	48.1%	Pred. No. 4.4e-10:	Length 973:
Matches 337:	Conservative	4:	Mismatches 346:
		Indels	14:
		Gaps	2:

QY	1107	CCGAAATACAGAGTGGCTCTACAGAGTGAACCTCTCGCTTCAAGGGGAGCTGGTCTT	1166
Db	270		
QY	1167	CCTGGAATGGCATGAGGACCCACCATCGCTGTGGCCCTCGTGGCCGCTCGGGCTTCT	1226
Db	330	CTTCTCTGGAAGGACACTTTTGGCGGTTCTATGGCCCTTCTCCACTCGTGGGATCCT	389
QY	1227	CAGCACCTTGCGCATCTCGTGATATTTTGGAGGACTTCCAGACCCATAGTTCCGTC	1286
Db	390	CTGTGCTCTGWAACATCGGCTCTGTTTCTGTACAAAGGGGACACTCGGAGTGAAGC	449
QY	1287	GGCTGGGGGGCCCATATGCTCTTCTGATGCTGTACACTGCTGTGGGGCATATGATGGT	1346
Db	450	AGCAGAGGGGACCCCTTACGCAAGCATCTCTTCCCTGGTCAATCACTACATCAGGC	509
QY	1347	CCCGGTGTACGTGGGGGGCCGCCAAGTCTCCACTGCTTCCGCGCCAGGCCCTCTTTC	1406
Db	510	CGCGCTGTGTTTGGGGCGAGCMCAGCAACCTTCAGTGAAGGCCGCGCAGGTGCTTTCGG	569
QY	1407	CGTGTGCTTTCACAAATTTGATCTCTCTGTATGAGCGGTGCTTTCACATCGTCTGGC	1466
Db	570	CATCAGCTTCACTCTGTGTATCTGTGATCTGTGATCTGTGTAAGACTTGGCAGATCTGTAGC	629
QY	1467	CTTCAAGATGGCCACGCCCTTCCACGGCCCTACAGTACTGTGGGTCCGCTACCAAGGGCC	1526
Db	630	CTTCAGTTTCAACCCCGGGGTGACGGAGCTGTGAGAA-----GATCTATCGGC	680
QY	1527	CTACGTCTCTATGGCATTTATCAGGGTACTAAAATGCTCATTTGGTAAATTTGGCATGCT	1586
Db	681	GTACGCCATTCATCACCGTCTCGTGGCCCTTCCAGCGGTGATCTGTACATCTGCTGGCTGT	740
QY	1587	GGCACGGCCTAAGTCCACCCCGGACTAGACCCGATAGCCCAAGATCACAATTTGTCTC	1646
Db	741	CCTCCAGAGCCCGTATCATACACCATCGGGCAGCCACACTGTTTTGAATGACTGCA	800
QY	1647	CTGTAAACCCCAATACCCGACAGCGCTGCTTCAACACAGCCCTGGACCTGCTGCTCTC	1706
Db	801	CGAAATCTGTACGTGGCCCTTTCGGGGTATCTGTGGGTATCATCGCCCTCTGGGCCCTTT-	859
QY	1707	AGTGTGGGTTCACCTTGGCTACATGAGGGAAGAGTGTGCCACCACTCAAGAGAGC	1766
Db	860	----GTGTGCTTCGTCTGCTTCCAAAGGGGCGCAAAATWACCGGAGCATGCAACGAGGC	915
QY	1767	CAAGTTCATCAACCTCAGCATGACTTCTATTTACCTCAT	1807
Db	916	CAAGTTCATCAACCTCAGCATGCTGCTTACTCTATTTCT	956

RESULT 6	
LOCUS	CNS0506r/c
DEFINITION	CNS0506R 1038 bp DNA linear GSS 26-JUL-2000
ACCESSION	Tetradon nigroviridis genome survey sequence T7 end of clone
VERSION	006M17 of library B from Tetradon nigroviridis, genomic survey
KEYWORDS	sequence.
SOURCE	AL315180
ORGANISM	AL315180.1 GI:9548068
	GSS: genome survey sequence.
	Tetradon nigroviridis.
	Tetradon nigroviridis
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
	Acanthomorphi; Acanthopterygii; Percomorphi; Tetraodontiformes;
	1 (bases 1 to 1038)
REFERENCE	Roset Crollius,H., Jalllon,O., Dasilva,C., Bouneau,L., Fisher,C.,
AUTHORS	Bernot,A., Fitzames,C., Wincker,P., Brottier,P., Quetier,F.,
	Saurin,W. and Weissensbach,J.
TITLE	Estimate of human gene number provided by genome-wide analysis
JOURNAL	using Tetradon nigroviridis DNA sequence
MEDLINE	Nat. Genet. 25 (2), 235-238 (2000)
PUBMED	20296633
	10835645

REFERENCE	AUTHORS	TITLE	JOURNAL	MEDLINE	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	SOURCE
2 (bases 1 to 1038)	Croillius,H.R., Jallion,O., Dasilva,C., Ozouf-Costaz,C., Fitzames,C., Fischer,C., Bouneau,L., Billault,A., Queller,F., Saurin,W., Bernot,A. and Weissenbach,J.	Characterization and repeat analysis of the compact genome of the freshwater pufferfish Tetraodon nigroviridis	Genome Res. 10 (7), 939-949 (2000)	20359637	10889143	3 (bases 1 to 1038)	Genoscope.	Direct Submission	Submitted (12-APR-2000)	This sequence is a single read and was generated as part of a large scale clone-and-sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at http://www.genoscope.cns.fr/Tetraodon .	Location/Qualifiers	1..1038

BASE COUNT	242 a	249 c	332 g	201 t	14 others
ORIGIN	/note="Genoscope sequence ID : COAB006AG09C1-end : T7"				

Query Match	4.89;	Score 96.4;	DB 17;	Length 1038;
Best Local Similarity	48.18;	Pred. No. 8.8e-10;		
Matches 378; Conservative	0;	Mismatches 391;	Indels 17;	Gaps 3;

[illegible]

Db	358	TGCTTTTGTTCCTTTATATGACAGCAGTGAGTGTACTGGCAGCGGTGGAGANGG	299
Oy	1871	TGTFACATGTGTCTCAACCTCCGCGCATACAGCTTGACTTGGCCCCAAGTCTACA	1930
Db	238	TGTCATCTCTATCTCCACACATACAGGCATCTCTGCGCTCACATTCTTTCCAAATCTTACA	239
Oy	1931	TGATCTCTCTTACACCGGAGCAGACACGCCGCCCTACTCTTCAACAGCATGATCCAGGGCT	1990
Db	238	TATATTCTTTCAAAGAAGACATAAATACCAGAGCGCGTTTCAGGAAGACTGTATAGAT	179
Oy	1991	ACACCA	1996
Db	178	ACTCCA	173
RESULT 7	A1742401/c	562 bp	mRNA linear EST 19-DEC-1999
LOCUS	A1742401	wg40e02.x1 Soares NSF_F8_9W_OT_PA_P.S1 Homo sapiens cDNA clone	
DEFINITION	IMAGE:2367578.3; similar to FR:093557.093557 PUTATIVE ODORANT		
RECEPTOR	A1742401	EST.	
VERSION	A1742401.1	GI:5110689	
KEYWORDS	human.		
SOURCE	Homo sapiens		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.		
TITLE	NCI-Cancer Institute, Cancer Genome Anatomy Project (CGAP),		
JOURNAL	Tumor Gene Index		
COMMENT	Unpublished (1997)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-ft@mail.nih.gov		
	This clone is available royalty-free through LNL; contact the		
	IMAGE Consortium (info@image.lnl.gov) for further information.		
	Insert Length: 840 Sd Error: 0.00		
	Seq primer: -400P from Glbco		
FEATURES	High quality sequence stop: 468.		
source	Location/Qualifiers		
	1..562		
	/organism="Homo sapiens"		
	/db_xref="taxon:3606"		
	/clone_image="IMAGE:2367578"		
	/clone_id="Soares_NSF_F8_9W_OT_PA_P.S1"		
	/lab_host="DH10B"		
	/note="Organ: pooled; Vector: pTR3D-Pac (Pharmacia) with		
	a modified polylinker; Site.1: Not I; Site.2: Eco RI;		
	Equal amounts of plasmid DNA from five normalized		
	libraries were mixed, and ss clones were made in vitro.		
	Following HAP purification, this DNA was used as tracer in		
	a subtractive hybridization reaction. The driver was		
	PCR-amplified cDNAs from pools of 5,000 clones made from		
	the same 5 libraries. The pools consisted of the following		
	libraries and clonesids: Soares NbHSF pool 1:		
	309384-310919, 323208-325895 Soares NbZHP pool 1:		
	145032-147335, 147720-148103, 148872-149255, 15002 -		
	150407, 151176-152327 Soares NbZHR8-9W pool 1:		
	758880-760583, 772104-774407 Soares NbHRA pool 1:		
	304776-306311, 320136-322823, 326280-326663 Soares NbHOT		
	pool 1: 723720-726407, 735080-740999 Subtraction by Bencio		
	Soares and M. Fatima Bonaldo."		
BASE COUNT	116 a	161 c	167 g
ORIGIN	117 t	1 others	
Query Match	4.8%	Score 95.6	DB 9;
Best Local Similarity	56.8%	Pred. No. 9.9e-10;	Length 562;
Matches 176;	Conservative 0;	Mismatches 134;	Indels 0;
Gaps 0;			
Oy	1699	CTGGCTTAGGAGGGGGTTCAGGCTTCCCAATAGGGGCAAGAAGAGCTGCCACCAATAC	1758
Db	499	CTCTCTTCATCAGTGCCTTGCCGTGACAGCTACCTGGGTAAGAGACTTCCAGAGAACTAC	440

Oy	1759	AACGAGGCAAGTCAATCACCCTCAGCATGACTTCTATTTCAACCTCAATCGGTCCCTC	1818
Db	439	AACGAGGCAAAATGTGTACCTTCACGCTGCTCTTCACTTCGTGTCTGGATCCGCTTC	380
Oy	1819	TGCACCTTCATGTCTGCTCAAGGAGGGGCTGTGTACCATGTGTGACCTCTTGGTCACT	1876
Db	379	TTTCAACAGGGCAACGCTCTACGAGGCGAACTACCTCCTCTGCGGCCAACTATGATGGCTGGG	320
Oy	1879	GTGTCAACCTCCTGCGCATCAGGCTTGAGCTTCTGGGCCCCAAGTCAATGATCTCC	1936
Db	319	CTGAGCAAGCTCGAGCAAGCGGCTTGCGGGGTATTTTCTGTGCTTAAGTGTACTGATCTCTC	260
Oy	1939	TTTCAACCGGAGCGCAACAGCGCCGCGCTTACTCAACAGCATGATCCAGGGCTACACCATG	1996
Db	259	TGCGCGCCAGACCTCAACAGACAGAGCACTTCCAGGCGCTCCATTCAGGACTACAGAGAG	200
Oy	1999	AGGAGGGACT 2008	
Db	199	CGCTGCGGCT 190	

RESULT	8
B1462667	
LOCUS	
DEFINITION	B1462667 938 bp mRNA EST 21-ANG-200
ACCESSION	603202107F1 NIH_MGC_97 Homo sapiens CDNA clone IMAGE:5268264 5'
VERSION	B1462667
KEYWORDS	B1462667.1 GI:15253323
SOURCE	EST.
ORGANISM	human.
REFERENCE	Homo sapiens:
AUTHORS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
JOURNAL	1 (bases 1 to 938)
COMMENT	NIH-MGC http://mgc.nci.nih.gov/. National Institutes of Health, Mammalian Gene Collection (MGC) Unpublished (1999) Contact: Robert Strausberg, Ph.D. Email: cga@bsf.remall.nih.gov Tissue Procurement: Miklos Palkovits, M.D., Ph.D. CDNA Library Preparation: Michael J. Brownstein (NHGRI), Shihaki Toshiyuki and Piero Carninci (RIKEN) cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLN) DNA sequencing by: Incyte Genomics, Inc. Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov plate: URAM1676 row: m column: 01 High quality sequence stop: 739.
FEATURES	
source	Location/Qualifiers 1..938 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="IMAGE:5268264" /clone_lib="NIH_MGC_97" /lab_host="DH10B" /note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site.1: BamHI; Site.2: SalI-XhoI (gtcgac); Oligo-dt primed using primer 5'-TTTTTTTGTGGTTTN-3' size selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."
BASE COUNT	186 a 260 c 294 g 198 t
ORIGIN	
Query Match	4.6% Score 92.8; DB 13; Length 938;
Best Local Similarity	61.7%; Pred. No. 4.8e-09;
Matches 148; Conservative	0; Mismatches 92; Indels 0; Gaps 0

QY 1 ATACACTACAGCGCCATCAGGATGATGGTGAGACAGATGGGCTTCCGGCTTTGGTG 60
Db 273 ATTTACTATGGCGCCAGCAGCGACAGCTCAGCCGTAAGCGGCACTATACCTCTTTCTTG 332
QY 61 CGTACACACCCAGCGCCGACCCACCTCGAGGCAATGATGACATATGCTGCATTC 120
Db 333 CGCACCAATCCCCAATAGACAAATGATCAGGATGGAGACCAATGGTGTCTCTCTGAGAAAGTTC 392
QY 121 CGCTGGAACCTGGATCATTTGTGCTGTGAGCAGCCAGACACTATGGCCGCGACAAATGGCC 180
Db 393 GGGGTGACCTGGATCTCTGTGGTGACGACATGACGACTATAGGGCACCTGAGGGGTGAG 452
QY 181 CTGCTGGCGAGCGGCTGGCCCGCGGACCAATCTCATTCGGCTTCCAGAGACGCTGGCC 240
Db 453 GCACCTGGAGAACCGGCCACTGTGTAGCGCATCTGCATTTGCTTTAAAGACATCATGGCC 512

RESULT 9	CNS03DLP/c	LOCUS	DEFINITION
	CNS03DLP	965 bp	DNA
			linear
			GSS 15-MAY-2000
			Tetradon nigroviridis genome survey sequence PUC-Orig end of clone
			016a11 of library G from Tetradon nigroviridis, genomic survey

ACCESSION	AL228534
VERSION	AL228534.1
KEYWORDS	GSS; genome survey sequence
SOURCE	Tetraodon nigroviridis.
ORGANISM	Tetraodon nigroviridis

REFERENCE	AUTHORS
1 (bases 1 to 965)	Roest-Crollius, H., Tallon, O., Dasilva, C., Bouneau, L., Fisher, C.,

TITLE	Human gene number estimate provided by genome wide analysis using
JOURNAL	Tetradon nigroviridis DNA sequence
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 965)
	Roest-Crollius, H., Jallion, O., Dasilva, C., Fitzames, C., Fisher, C.,

TITLE	Characterization and repeat analysis of the compact genome of the freshwater pufferfish <i>Tetraodon nigroviridis</i>
JOURNAL	Unpublished
REFERENCE	3 (bases 1 to 965)

COMMENT This sequence is a single read and was generated as part of a large-scale clone-end sequencing project of the Tetraodon nigroviridis genome. For more information, please take a look at <http://www.genoscope.cns.fr/tetraodon>.

FEATURES	SOURCE	LOCATION/QUALIFIERS
		1..965
		/organism="Tetraodon nigroviridis"
		/db_xref="taxon:99883"
		/clone="016a11"
		/clone_lib="G"
		/note="Genoscope sequence ID : C0Bc016A06Sp1-end : PUC-01.1"
BASE COUNT	317 a	169 c 225 g 249 t 5 others
ORIGIN		

Query Match	4.5%;	Score 90.2;	DB 117	Length 965;
Best Local Similarly	55.6%;	Pred. No. 1.7e-08;		
Matches 170;	Conservative 1;	Mismatches 135;	Indels 0;	Gaps 0;

Qy 1691 TGGACCTGCTGCCTCAGTGGTGCGGTTCAGCTCCGCCATCAATGGGCAAGAAGACTGCCGA 1756
|| || || | | | | | | | | | | | | | | | |
Db 925 TGCATTCTAATGTCTGTGCGCATCTCTGCTTTACTTTCCTCTACATGGGGAAAGMCTCCCGA 866

QY	1751	CAACTCAACAGAGAGCCCAATTCATACACCTGATGACACCTTCATTTTCAACCTCATCG	1810
Db	865	AGAACTCAACGAGAGCCCAAGACCAATACCTTCGTCTTGCTTGTGATCTCCACCTGGA	806
QY	1811	TCCTCCCTCGACACCTTCATGCTTCGCTTAACAGCGGGGTGCTGGTCAACATCGTGACCTCT	1870
Db	805	TCATATTTTNGCCCGCTTCATGCTTACACGGCAAAATCAATCCACACCTTCACAGCCC	746
QY	1871	TGGTACGTCGTCCAAACCTCTGCGCATCAGCTGGGGCTACTTCGGGCCCAAGTGTCTCA	1930
Db	745	TGGCGGTGCTCTCCAGAGGCTTACTGCTTCTCTGTGGTAATTTCTCCCAATATTTTCA	686
QY	1931	TGATCTCTTCTTACCCGAGACGCAACAGCCGCCCTACTTCAACAGCATGATCAGGGCT	1990
Db	685	TTATCATCTTTCAACCTCCACAGAACACACCCAGAAATCTTCCAAAGGTCTGATTGAGACT	626
QY	1991	ACACCA	1996
Db	625	ACACCA	620

RESULT 10	LOCUS	DEFINITION
AI562167/c	AI562167	298 bp mRNA linear EST 25-MAR-1999
		vw/3gd10.x1 Stratiogene mouse heart (49373716) Mus musculus cdna clone

ACCESSION	AI562167	GI:4513512
VERSION	AI562167.1	
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

REFERENCE
AUTHORS
Marina, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
1 (bases 1 to 298)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 298)

TITLE	The WashU-NCI Mouse EST Project 1999
JOURNAL	Unpublished (1999)
COMMENT	Contact: Marra W/WashU-NCI Mouse EST Project 1999

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:663147
This clone was previously sequenced on the 5' end only, this new
data is from the 3' end
Possible reversed clone: similarity on wrong strand
Possible reversed clone: polyA not found
High quality sequence stop: 286.

```

FEATURES
Source
Location/Qualifiers
1..298
/organism="Mus musculus"
/strain="NIH Swiss"
/db_xref="taxon:10090"
/clone="IMAGE:1260595"
/clone_lib="Stratagene mouse heart (#937316)"
/sex="pooled"
/tissue_type="heart"
/dev_stage="13 day embryos"
/lab_host="SOLR (kanamycin resistant)"
note="Organ: heart; Vector: pBluescript SK-; Site:
EcoRI; Site_2: XhoI; Cloned unidirectionally. Prim:
Oligo dT. 93 pooled NIH/Swiss 13 day embryo hearts
Average insert size: 1.0 kb. Uni-ZAP XR Vector: -5'
adaptor sequence: 5' GAATTCGCGACGAG 3'-3' adaptor
sequence: 5' CTCGACGATTTTTTTTTTTTTT 3'"

```

BASE COUNT	59 a	101 c	87 g	51 t	ORIGIN
Query Match					
Best Local Similarity	59.2%	4.1%	Score 82.8;	DB 9;	Length 298;
Matches 141;	Conservative	0;	Mismatches 97;	Indels	0;
					Gaps 0;
QY	2	TCACCTACAGGGCCATCAAGCGATAGCTGCGGACAGACAAGTGGCGTTCCCGGCTTTGCTGC	61		
DB	291	TCAGCTTAAGAGCGCGACGCGCGATCTCAGGAGGAGGCAAGTCCCGCTCTCTCTGC	232		
QY	62	GTACACACACCCAGCGCCACACACACACGTCGAGGGCCATGGTGCAGCTGATCTGCACCTTCC	121		
DB	231	GCACCATCCCGACGCGATTAAGTACACGAGTGGAAGTATAGTGGGGGTGTGTGCAGAGCTTGC	172		
QY	122	GCTGGAACCTGGATCACTTTGTGCTGGTGGACAGCGACACCTATGCGGCCGCGACAAATGGCCACG	181		
DB	171	GCTGGGCTGTGATTCGCGCTCGTGGCACACTATGATGACTACGCGGACGCTGGGCGCTACAGG	112		
QY	182	TGCTTGGGCGAGCGGCTGGCGCCGCGCGACATGTGATGCGCTTCCAGAGACGCTGGCC	239		
DB	111	CGCTGGAGGAGCTGGCCCACTCCACGCGGGCATCTGTGCGCTTTCAGAGACGTGGTGC	54		

RESULT 11	
BB618551/c	
LOCUS	BB618551 666 bp mRNA linear EST 26-OCT-2001
DEFINITION	BB618551 RIKEN full-length enriched, 8 days embryo Mus musculus
	CDNA clone 5730408M21 5', mRNA sequence.

ACCESSION	BB618551	GI:16458094
VERSION	BB618551.1	
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

REFERENCE
AUTHORS
Arakawa, T., Carrincci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A.,
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Euthalia; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 666)

TITLE
JOURNAL
COMMENT

Himemoto,O., Horii,F., Ishiy,I., Ito,M., Kawai,T., Konno,H., Kouda,
,M., Koysa,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M.,
Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasak
,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H.,
Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T.,
Muramatsu,M. and Hayashizaki,I.
RIKEN Mouse ESIS (Atakawa,T., et al. 2001)
Unpublished (2001)
Contact: yoshihide hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gs.c.riken.go.jp,
URL: <http://genome.gsc.riken.go.jp/>
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.,
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. *Genome Res.*, 10 (10), 1617-1630 (2000)
wagii, K., Fujiwara, S., Inoue, K., Togawa, Y., Itawa, M., Ohara, E., Watabiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.,
RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multicapillary sequencer. *Genome Res.* 10 (11), 1757-1771 (2000)
Kono, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.,
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. *Genome Res.* 11 (2), 281-289 (2001)
Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamakawa, I., Alzawa,

K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.
Computational Analysis of Full-length Mouse cDNAs Compared with Human Genome Sequences. *Mamm. Genome*. 12, 673-677 (2001).
Please visit our web site (<http://genome.ysc.riken.go.jp>) for further details.
e mouse tissues.
Location/Qualifiers
I. .666

```

/organism="Mus musculus"
/strain="C57BL/6J"
/db xref="taxon:10090"
/clone="5730408W1"
/clone_lib="RIKEN full-length enriched, 8 days embryo"
/sex="mixed"
/dev_stage="8 days embryo"
/lab_host="DH10B"
/notes="Site 1: SalI; Site 2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory 1
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA w
primed with a primer [5'
GAGAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTCTT 3'], cDNA w
GAGAGAGAGAGATCCAGAGCTCTTTTCTTTTCTTTTCTT 3'}. cDNA w
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length
cap-trapper. cDNA went through one round of subtraction
Rot = 100.0 Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGAGATTCGAGATTAAATTAATTCGCCCCCCCCC 3'}. cDNA
was cloned into the XhoI and BamHI sites. Vector: a
modified pluscript KS(+) after bulk excision from Lam
FLC I. Cloning sites, 5' end: SalI, 3' end: BamHI."

```

Query Match	4.1%	Score 82.2	DB 10	Length 666
Best Local Similarity	54.0%	Pred. No. 6.8e-07		
Matches 168	Conservative 0	Mismatches 143	Indels 0	Gaps
Qy 1697 TGGTGCCTCAGTGGTGGGTTTCACCTTGGCTCATGGGCAAGAGCTGCCACCACT	1756			
Db 571 TCTCTCTTCACACACACCTTTGCTGTGACGTACCTGGGTAAGAACCTGGGGAACCT	512			
Qy 1757 ACAAGAGGCGCAATTTCACCTCAGACATGACCTCTATTTCACCTCATCGCTGCC	1816			
Db 511 ATAAAGAGCAAAATGTGTACCTTACACCTGTGCTCTCCACTTGGATTCTGGATGCTT	452			
Qy 1817 TCTGCACCTCATGTGTGCTTACAGCGGGGTGCTGTACCATGTGACCTTTGGTGA	1876			
Db 451 TCTTTCACCAATGTCTAAGCATTTTACAGGCGACGTACTACCGCGGTCAATGTCTGGCAG	392			
Qy 1877 CTGTGCTCAACCTCCCTGGGCGATACACCTGGGCTACTTGGGCCCAATGCTAATATATCC	1936			
Db 391 GGGTGGCACTGTGAGTGGCGGCTTCAAGCGCTATTCTCCCTAAATTTGTAGTACTTC	332			
Qy 1937 TCTTTTACCCGGAGGCGCAACAGCGCGGCTCTACTTCAACAGATGATTCAGGGGTACAGCA	1996			
Db 331 TCTGCGGCTCAGAACTCAACAACAAGAACTTTACAGGCTTCATCCATCCAGACTACAGCA	272			
Qy 1997 TGAAGAGGGAC 2007				
Db 271 GGGCGTGGCGC 261				

RESULT 12	
AZ535744	
LOCUS	AZ535744
DEFINITION	816 bp DNA linear GSS-03-NOV-2-
ACCESION	Entamoeba histolytica Sheared DNA Entamoeba histolytica genomic, DNA sequence.
VERSION	AZ535744.1 GI:11092691

ORGANISM

Mus musculus

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 577)

AUTHORS

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamill,C., Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A. and Wright,D., Weiss,R.

TITLE

Mouse whole genome scaffolding with paired end reads from 10kb

JOURNAL

Plasmid inserts
unpublished (2000)

COMMENT

Contact: Robert B. Weiss

University of Utah Genome Center

Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT 84112, USA

Tel: 801 585 5606

Fax: 801 585 7177

Email: ddunne@genetics.utah.edu

Insert length: 10000 Std Error: 0.00

Plate: 0076 row: F column: 07

Seq primer: CACACAGCAACACGCTATGACC

Class: plasmid ends

High quality sequence stop: 577.

Location/Qualifiers

FEATURES

source

1..577
/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/clone="UUCG1M0076F07"

/clone_1lb="Mouse 10kb plasmid UUCG1M library"

/sex="Male"

/lab_host="E. coli strain XL10-Gold, T1-resistant, F-"

/note="Vector: PMD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson

Laboratory Mouse DNA Resource

(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a

0.005 inch orifice at constant velocity. The sheared DNA

was blunt end-repaired with T4 DNA polymerase and T4

polynucleotide kinase. Adaptor oligonucleotides were

ligated to the blunt ends in high molar excess. The

adapted DNA was purified and size-selected for a 9.5 to

10.5 kb range using preparative agarose gel

electrophoresis. Vector DNA was prepared from a derivative

of pMD42 (g1473214|gblaf129072.1), a copy-number

inducible derivative of plasmid R1. The vector was ligated

with adaptors complementary to the insert adaptors and

purified. The sheared, adapted mouse DNA was annealed to

adapted vector DNA, and transformed into

chemically-competent E. coli XL10-Gold (Stratagene) cells

and selected for ampicillin resistance."

BASE COUNT 144 a 142 c 178 g 113 t

ORIGIN

Query Match

Best Local Similarity 3.8%; Score 76.4; DB 17; Length 577;

Matches 95; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

DB 472 GGTCCCATATCCATGTTCTTAAGAGTGTGCCAGCTGGCAAAATGAACCAATAGG 413

OY 981 GATCCCTATGTCATGTCCTCAAGAGTGCATGAGGCAAAAGAGAGCCTGTGGG 1040

DB 472 GGTCCCATATCCATGTTCTTAAGAGTGTGCCAGCTGGCAAAATGAACCAATAGG 413

OY 1041 CATCCAGCTGCTGCTGAGTGCATGAGCTTCCCGGCACTTCCCTCAACACAC 1100

DB 412 CCTCCACCCGTCGTCCTTGAGTGTGGACTGTCCGCGGCACTTCACTCAACCGATC 353

OY 1101 TGAATG 1106

DB 352 AGTAGG 347

OY 1101 TGAATG 1106

DB 352 AGTAGG 347

OY 1101 TGAATG 1106

DB 352 AGTAGG 347

OY 1101 TGAATG 1106

DB 352 AGTAGG 347

OY 1101 TGAATG 1106

DB 352 AGTAGG 347

OY 1101 TGAATG 1106

DB 352 AGTAGG 347

OY 1101 TGAATG 1106

DB 352 AGTAGG 347

OY 1101 TGAATG 1106

DB 352 AGTAGG 347

Job time : 2843.62 secs

THIS PAGE BLANK (USPTO)

Db	241	ACACGCGACGCCCAACGAAACATGACGTACAGAGACGCCAGCGCTGGTACATTGTG	300
QY	301	GACAAAGTGCAGCAGACACACAGCGCGCGTGTGTGTCTCTCGCCGACCTGCACCTG	360
Db	301	GACAAGGTGCAGCAGACACACAGCGCGCGTGTGTGTCTCTCGCCGACCTGCACCTG	360
QY	361	TACACCTCTTCATGAGGGGTGCGCGCAGAACTTTCACGGCGCGCTGTGGATTCGCTTC	420
Db	361	TACACCTCTTCATGAGGGGTGCGCGCAGAACTTTCACGGCGCGCTGTGGATTCGCTTC	420
QY	421	GAGTCTGGGCGCATTCGACCGCGTCTTCACAACCTCAGAGAGTGGGCGACCTTGGGAC	480
Db	421	GAGTCTGGGCGCATTCGACCGCGTCTTCACAACCTCAGAGAGTGGGCGACCTTGGGAC	480
QY	481	TTTCTGGGCGATTCACATTCACAGCGGTGCCATCCGGGCTTCAGTAGTTCGCGCAGTGG	540
Db	481	TTTCTGGGCGCATTCACATTCACAGCGGTGCCATCCGGGCTTCAGTAGTTCGCGCAGTGG	540
QY	541	GGCCCAACAGCGTGGGGCGGCCACCCCTTCAGAGAGACCAAGCAAGATATATCTGCACACAG	600
Db	541	GGCCCAACAGCGTGGGGCGGCCACCCCTTCAGAGAGACCAAGCAAGATATATCTGCACACAG	600
QY	601	GAGTGCAGCACTCTCTGAAGCGCACCTTCTTCCTTAACACCAATTCACAGGCTCTCGAG	660
Db	601	GAGTGCAGCACTCTCTGAAGCGCACCTTCTTCCTTAACACCAATTCACAGGCTCTCTGG	660
QY	661	GAGCGTGTCTACACAGCGGTACTCTGCGGTATGTCTGTGGCCCATCCCTGCACAGC	720
Db	661	GAGCGTGTCTACACAGCGGTACTCTGCGGTATGTCTGTGGCCCATCCCTGCACAGC	720
QY	721	CTTCTCGGCTGTGACAAAGACACTGTGACCAAGAGGTGTCTACCCCTGTGCGACTGT	780
Db	721	CTTCTCGGCTGTGACAAAGACACTGTGACCAAGAGGTGTCTACCCCTGTGCGACTGT	780
QY	781	GAGGAGTCTTGAAGGTCACTTACCTCTGCTGAGACACCAAACTCTTTCGACCCGAA	840
Db	781	GAGGAGTCTTGAAGGTCACTTACCTCTGCTGAGACACCAAACTCTTTCGACCCGAA	840
QY	841	GGGAGCGTGGCTGTGCACTTGGAGATTGTCCAACTGGCAATGGGACCGGACCAAGATCCC	900
Db	841	GGGAGCGTGGCTGTGCACTTGGAGATTGTCCAACTGGCAATGGGACCGGACCAAGATCCC	900
QY	901	TTTCCAGACCGTGGCTCTTACTACCCCTGTGACGACAGCTGAAGAACTCAAGACTCT	960
Db	901	TTTCCAGACCGTGGCTCTTACTACCCCTGTGACGACAGCTGAAGAACTCAAGACTCT	960
QY	961	CTGCACACCGTCAACACAGACAGATCCCTATGTCCATGTCATGTTCCAAAGGTGCCAGTACAGG	1020
Db	961	CTGCACACCGTCAACACAGACAGATCCCTATGTCCATGTCATGTTCCAAAGGTGCCAGTACAGG	1020
QY	1021	CAAAAAGAAAGCCTGTGGGCATTCACGTCGTCTGTGAGTGCATGACGTCCCTCC	1080
Db	1021	CAAAAAGAAAGCCTGTGGGCATTCACGTCGTCTGTGAGTGCATGACGTCCCTCC	1080
QY	1081	GGCACCTTCTCAACACCACTGAATGCCGAATTAACAGATGTCTTACCAAGATAGAC	1140
Db	1081	GGCACCTTCTCAACACCACTGAATGCCGAATTAACAGATGTCTTACCAAGATAGAC	1140
QY	1141	TCTCGCTTCAAGGGCGAGGTCTTCCTCGGAAATGGCATAGAGCACCACTTCCTGTG	1200
Db	1141	TCTCGCTTCAAGGGCGAGGTCTTCCTCGGAAATGGCATAGAGCACCACTTCCTGTG	1200
QY	1201	GCCCTGCGGCGCGCGCGGCTCTTCACAGACCTTGCCATCGTGATATTCGAGAG	1260
Db	1201	GCCCTGCGGCGCGCGCGGCTCTTCACAGACCTTGCCATCGTGATATTCGAGAG	1260
QY	1261	CACCTTCAGACACCCATATGTTGCTCGTGCGGTGGGGGCCCATGAGCTTCTGATGTGACA	1320
Db	1261	CACCTTCAGACACCCATATGTTGCTCGTGCGGTGGGGGCCCATGAGCTTCTGATGTGACA	1320
QY	1321	CTGCTGTGCTGGATATCATGTGTGCTCGGTGTACGTGGGCGGCCAAAGTCTTCAC	1380
Db	1321	CTGCTGTGCTGGATATCATGTGTGCTCGGTGTACGTGGGCGGCCAAAGTCTTCAC	1380

Db	1321	CTGCTGCTGGTGGCATACATGATGGTGTCCGGGTCTACGTGGGGGCGCGCAAGGTCTCCACG	1380
Qy	1381	TGCGCTCTGCGCCGACAGGCCCTCTTTCCTCTGCTTCACAAATTTGGCATCTCTGTATGGCC	1440
Db	1381	TGCGCTCTGCGCCGACAGGCCCTCTTTCCTCTGCTTCACAAATTTGGCATCTCTGTATGGCC	1440
Qy	1441	GTGGGTTCTTTCGAGATCGTCTGGGCTTCGAATATGGGAGCGGCTTCCACAGCGCCCTAC	1500
Db	1441	GTGGGTTCTTTCGAGATCGTCTGGGCTTCGAATATGGGAGCGGCTTCCACAGCGCCCTAC	1500
Qy	1501	AGCTACTGGGTCGCGCTACGAGGGGGCCCTACGTCCTATGGCATTTTATACGGTACTACAA	1560
Db	1501	AGCTACTGGGTCGCGCTACGAGGGGGCCCTACGTCCTATGGCATTTTATACGGTACTACAA	1560
Qy	1561	ATGGTCATTTGGTGAATTTGGCATGTGGCACGGGCTCAGTCCACCCCGTACTGACCCC	1620
Db	1561	ATGGTCATTTGGTGAATTTGGCATGTGGCACGGGCTCAGTCCACCCCGTACTGACCCC	1620
Qy	1621	GATACCCCAAGATACAAATTTGCTCTGTAACCCCACTACCGCAACAGCCTGCTGTTTC	1680
Db	1621	GATACCCCAAGATACAAATTTGCTCTGTAACCCCACTACCGCAACAGCCTGCTGTTTC	1680
Qy	1681	AACACACAGCCTGAGCCTGCTCTCTGCTCAGTGGTGGTTTCAGTTCGGCTACATGGGCAAA	1740
Db	1681	AACACACAGCCTGAGCCTGCTCTCTGCTCAGTGGTGGTTTCAGTTCGGCTACATGGGCAAA	1740
Qy	1741	GAGTGCCACCAACATACAAAGAGGCCAAGTTTCATCAGCTTCAGCATGCACTCTATTTC	1800
Db	1741	GAGTGCCACCAACATACAAAGAGGCCAAGTTTCATCAGCTTCAGCATGCACTCTATTTC	1800
Qy	1801	ACCTCATCCGCTTCCCTCTGACCTTCATATCTGCTTAAGAGGGGGTGTGGTACACATC	1860
Db	1801	ACCTCATCCGCTTCCCTCTGACCTTCATATCTGCTTAAGAGGGGGTGTGGTACACATC	1860
Qy	1861	GTGACCTCTTGGTCACTGTGCTCAACCTCTGCGCATCAGCCTGGGCTACTTTCGGCCC	1920
Db	1861	GTGACCTCTTGGTCACTGTGCTCAACCTCTGCGCATCAGCCTGGGCTACTTTCGGCCC	1920
Qy	1921	AAGTGTACATGATCTCTTCTACCCGGAGCGCAACAGCCCGGCTTACTTCAACAGCATG	1980
Db	1921	AAGTGTACATGATCTCTTCTACCCGGAGCGCAACAGCCCGGCTTACTTCAACAGCATG	1980
Qy	1981	ATCCAGGCTTACACCATGAGGAGGACTAG	2040
Db	1981	ATCCAGGCTTACACCATGAGGAGGACTAG	2040

OTHER INFORMATION: rat G-protein coupled receptor (GPCR) B4

OTHER INFORMATION: nucleotide sequence
US-09-361-631-3

Query Match	61.3%;	Score 1232.8;	DB 4;	Length 2993;
Best Local Similarity	76.9%;	Pred. No. 1.9e-25;		
Matches 1568;	Conservative	0;	Mismatches 442;	Indels 30; Gaps 4

QY	1	ATACCCATACAGGCGCATACAGATGAGACGACGAGCAAGAGTGGCTCCCGGCTTTGGCTG	60
Db	505	ATCACATACAGCGCCATCTCTCGAACAAGCTGGGGAACAAGCGACTCTCTAGACATGCTA	564
QY	61	CGTACACACCCAGCGCCGACCAACAGTCGAGGCGCATGGTCAAGTATGCTGCATTC	120
Db	565	CGCACATATGCCACAGCGCCACCCACCAATCGAGGCCATGATGATGGTTACATTC	624
QY	121	CGCTGGAACATGGAATCATTTGGTGTGGTGAAGAGCACTATATGGCCCGACAAATATGCCAG	180
Db	625	CAATGGAACTGGAGTTGTGGTGTGGTGAAGAGCAAGATATACGGCCGCGAAGAACGCCAC	684
QY	181	CTGCTTGGCAGCGCGTGGCC---CGGCGCGAATCTGCATCGCCTTCCAGAGACAGCTG	237
Db	685	CTGTGTAGGACAGGCTGTGACCAAAAGAGGAGCATTCGATTGGCTTCCAGGAGGTCTG	744
QY	238	CCCACTATGAGCGCCCAACAGAAATGAGTGAAGAGAGCGCAGCGCCTGGTACCAATT	297
Db	745	CCCATACCTTAGTTCACACCCAGTCAATGAGGTCCGAGAGCACAGAACATGTGACAAATC	804
QY	298	GTGACATCAGTGTGAGAGACAGACAGCGCGCTGTGTGTGTCTTCTCGCCGACCTGCAC	357
Db	805	CTGGACAAAGTGGCGGGAGCTGGGGCGCGCTGTGTGTGTCTCTCGCCGAGCTGAGC	864
QY	358	CTGTACCACTTCTTCAATGAGTGTGGCGCCAGAACTTACGGGCGCGGTGTGATGCC	417
Db	865	CTGTATATGCTTCTTTCACGAGGTGCTCCGCTGGAACTTACGGGTTTGTGTGATTCGCC	924
QY	418	TCCAGATCTGGGGCATACGACCCGGTCTGACAACTCAGGAGAGTGGGCGACATGGGG	477
Db	925	TCTGAGTCTGGGGTATCGAACCCAGTTCTGTCATACCTACGGAAGCTGGCCACAGGGT	984
QY	478	ACCTTCTGGGGCATACACATCATCAGAGCGTGGCCATCCGGGGTTCAGTAGATTTCCGCGAG	537
Db	985	ACTTTTCTGGGGCATACATTCACAGAGGGTGTCCATCCCTGGGCTTACGTACAGTTCCGAGTG	1044
QY	538	TGGGGCCCAACAGGCGTGGGGCGCCACCCCTCAGAGGAGACGAGCAGATACCTCGGCAAC	597
Db	1045	CGCGGTACAAAGCCAGGGATTCCCGTGGCTTACACAGACCAACCTGGGAGCGACTGCAAC	1104
QY	598	CAGGAGTGCACAACACTGCTGAACGGCACACTTGTCTTTCACACCAATTCACAGGCTCTCT	657
Db	1105	CAGAGCTGATGACCGCTGCTTGAACACACCAAGATCCTTACACAACTCCTTATCTTTCG	1164
QY	658	GGGGAGGCTGTGCTACAGCGGTACCTGTGGGGTATATGCTGTGGCCCATATCCCTCGAC	717
Db	1165	GGGGAGGCGCTGTGTACAGCGGTACTGTGGGAGTTTACGGCGGTGGCCCATCCCTCGAC	1224
QY	718	AGCTCTCTCGCTGTGACAAACAACACTCACCACAAAGGGTGTGTACCCCTGGAGAGTG	777
Db	1225	AGAATCTCGGGCTGTAAACGGGGTCCGCTGCACCAACAAAGGTCTACCCCTGGAGAGCTA	1288
QY	778	CTTGAAGAGATCTGGAAGGTCAACTTTCACCTCTCCGTGACACCAATCTTCTTGACGCG	837
Db	1285	CTCAGGAGATCTGGACACGTCACACTTCAGCGTCTCTGGGTACACCGGCTCTTCTTGAACCA	1344
QY	838	CAAGGGAGCTGGCTCTGCACCTTGGAGAGATTGTCCATATGGGCAATGGGACCGGAGCCAGAT	897
Db	1345	CAAGGGAGCAATGCGAGTCTTTTGGACATATCCCACTGGACATGGGAGCTTGAGCCAGAT	1404
QY	898	CCCTTCAGAGCGCTCGCTCTCTACTACCCCTGTCAGACGACACTCTGAAGAAT---CAAG	954
Db	1405	CCCTTCAAAGCATCGCTCTCTATTTCTCCACACGAAAGGCTTAACCTTAACAAAT	1466
QY	955	ACATCTCTGCACACGTCACAAACAACGATCCCATGTCCATGTCTTCCAAAGAGTGGCAG	1014

[illegible]

Query Match	9.3%	Score 187	DB 4	length 4134
Best Local Similarity	51.2%	Pred. No. 2.7e 31		
Matches 464	Conservative	0	Mismatches 440	Indels 3
				Gaps 1

QY 1099 ACTGATGCCCCGAAATACGAGTGTGCTACAGAGTGAGACCTCTGCTTCAAGCGGAG 1158
 Db 2209 ACAAGATGCCCCGAAATGATTTTGTGCGAATGAAACACAGCTGTGATGCCAAGAG 2268
 QY 1159 CTGGCTTCCGGAATGAGCATGAGGACACCAACATCCGCTGTGGCCCTCTGCGCCCTG 1218
 Db 2289 ATCGAGTACCTGTGCTGAGCGAGCCCTTGGAGATGCTGTGACCATTTTGCGCTACTG 2328
 QY 1219 GCGTTCCTCAGACACCTTGGCCATCTGTGTATTTCTGAGAGCACTTCCAGACACCATTA 1278
 Db 2329 GGCATTCCTGATCACCCTCTCTGTGTGGGGCTTTCATCAGATTGAGAACATCTCCATC 2388
 QY 1279 GTTGGCTCGCTGGGGGCCCATCTGCTCTGATGCTGACACTGCTGCTGGTGACATAC 1338
 Db 2389 GTGAGGCGACCAACCGGAGAGTGTGCTACCTGCTGCTCTTCTCCATCATCTGCTGCTT 2448
 QY 1339 ATGGTGTCCGGGTGTAGTGTGGGGGCCCAAGGTCTCCACTGCTGCTGCGCCAGGCC 1398
 Db 2449 TCCAGCTTGCATCTTATGTATGGGAGGAGCCAGGAGCTGACCTGTGCGCCAGCCG 2508
 QY 1399 CTCTTCCCTCTGCTTACAAATTTGCAATCTCTGTATGCGGTCGCTTTTCCAGATC 1458
 Db 2509 GCGTTTGGCATCAGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 2568
 QY 1459 GTCTGCGCTTCAAGATGAGCCAGCCGCTTCCAGCGCTTACAGCTACTGCGGCTGCTAC 1518
 Db 2569 CTGCTGTGCTTGTGA---GGCCAGATGCTCCACAGCCCTCCACCGCAAGTGGGCGCTC 2625
 QY 1519 CAGGGGCGCTTGTGCTGTATGAGCATTTTACAGGAGTCTCAAGTCAAGTGTATGTGTAT 1578
 Db 2626 AACCTGCAAGTCTCTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 2685
 QY 1579 GGCATGCTGAGCGAGCCCTCAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1638
 Db 2686 TGCTGTACACCGGCGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 2745
 QY 1639 ATGTCTCTGTATACCCCACTACCGCAAGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1698
 Db 2746 TTCACTACCTGTGAGAGGCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 2805
 QY 1699 CTGCTGTGAGTGTGAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1758
 Db 2806 CTCTCTGCGCGCATCTGCTTCTTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 2865
 QY 1759 AACGAGCGCAATTCATACCCCTCAGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1818
 Db 2866 AACGAGCGTAGTGTATCATCTTCAAGATGTTGATCTTCTTCAATGCTGTGATCTGCTTTC 2925
 QY 1819 TGCACCTTCAATGTCTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 1878
 Db 2926 ATCCCGCGCTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCT 2985
 QY 1879 GTGCTCAACCTCTGTGCGCATACAGCTGTGAGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 1938
 Db 2986 CTGCGCTTCCAGCTTGTGCGCTGTGCGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 3045
 QY 1939 TTCTACCCGAGCGCAACAGCCGCTTACTTCAACAGATGAGGCTTACACCATG 1998
 Db 3046 TTCAAGCGCGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGT 3105
 QY 1999 AGGAGCG 2005
 Db 3106 AAGGTGG 3112

RESULT 6
 US-08-485-588-1
 ; Sequence 1, Application US/08485588
 ; Patent No. 5688938
 ; GENERAL INFORMATION:
 ; APPLICANT: Edward M. Brown
 ; APPLICANT: Steven C. Hebert
 ; APPLICANT: Forrest H. Fuller

; APPLICANT: James E. Garrett, Jr.
 ; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 ; MOLECULES
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lyon & Lyon
 ; STREET: First Interstate World Center
 ; STREET: Suite 4700
 ; STREET: 633 West Fifth Street
 ; CITY: Los Angeles
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 90071
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: FASTSEQ
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/485,588
 ; FILING DATE: 7 June, 1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; PRIOR APPLICATION DATA: Including application
 ; PRIOR APPLICATION DATA: described below: 9
 ; APPLICATION NUMBER: 08/353,784
 ; FILING DATE: 9 December, 1994
 ; APPLICATION NUMBER: PCT/US/94/12117
 ; FILING DATE: 21 October, 1994
 ; APPLICATION NUMBER: U.S. 08/292,827
 ; FILING DATE: 23 August, 1994
 ; APPLICATION NUMBER: U.S. 08/141,248
 ; FILING DATE: 22 October, 1993
 ; APPLICATION NUMBER: U.S. 08/009,389
 ; FILING DATE: 23 February, 1993
 ; APPLICATION NUMBER: U.S. 08/017,127
 ; FILING DATE: 12 February, 1993
 ; APPLICATION NUMBER: U.S. 07/934,161
 ; FILING DATE: 21 August, 1992
 ; APPLICATION NUMBER: U.S. 07/834,044
 ; FILING DATE: 11 February, 1992
 ; APPLICATION NUMBER: U.S. 07/749,451
 ; FILING DATE: 23 August, 1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hebert, Sheldon O.
 ; REGISTRATION NUMBER: 38,179
 ; REFERENCE/DOCKET NUMBER: 213/005
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (213) 489-1600
 ; TELEFAX: (213) 955-0440
 ; TELEX: 67-3510
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 5275 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA to mRNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 515..3769
 ; OTHER INFORMATION:
 ; US-08-485-588-1

Query Match 8.63; Score 172.6; DB 1; Length 5275;
 Best Local Similarity 50.28; Pred. No. 3.4e-28;
 Matches 455; Conservative 0; Mismatches 449; Indels 3; Gaps 1;

QY 1101 TGAATGCCCGCAATACAGTGTGCTTACAGAGAGGAGCCTCTGCTTCAAGGCGGAGCT 1160
 Db 2266 TAAATGCCCGCAATACAGTGTGCTTACAGAGAGGAGCCTCTGCTTCAAGGAGGAT 2325
 QY 1161 GGTCTTCTGGAATGAGCATGAGGACCCACCATGCTGTGTGCTGTGCTGTGCTGTGCTGTG 1220


```

Db 2326 CGAGTTTCTGTCGTGGACCGCCCTTGGGATGCACTACAGCTTTGGTGTGGG 2385
QY 1221 CTTCCTAGACACCTGGCCATCTGGTGATATTTGGAGCAGCTTCCAGACCCATAGT 1280
Db 2386 CATTTCTCTACACCTTCCTGGTGGGCGTCTTCATCAAGTTCCGCAACAGCCCATCGT 2445
QY 1281 TCGGTGGGCGGGGGCCCATGCTTCCATGCTGACACTGACGCTGGGCGGATCATCAT 1340
Db 2446 CAAAGGACCAACCGGAGGCTCTCTATCTCTTCTCTCTCTCTCTCTCTCTCTCTCT 2505
QY 1341 GGTGGTCCCGGTGATGAGGCGCCGCAAGTCTCCACCTGCTCCGCGCCAGGCGCT 1400
Db 2506 CAGTCCCTGTTCTTCACTGGGAGAGCCCGAGAGTGAACGTCGCCCGCCAGCCGCG 2565
QY 1401 CTTCCTCTGCTTCAATTTGATCTCTGATGCGCGCTGCTTCTTCCAGATCGT 1460
Db 2566 CTTTGGCATCACTTCGCTGCTGATCTGCTGATCTGCTGATAAACCATCGGCTCT 2625
QY 1461 CTGGGCTTCAAGATGGCCAGCCGCTTCCACGCGCTACAGCTACGTCAGTCCGCTACA 1520
Db 2626 CCTGGTGTGTA---GGCCAAAGATTCACACAGCTTCCACCGGAGTGGGCGCTCAA 2682
QY 1521 GGGGCGCTACGCTCTCTATGCGATTTATCAGCGTACTCAAAATGGCATTTGTAATTGG 1580
Db 2683 CCTGCAATCTGCTGCTGCTGCTCTCTGACCTTCACTGACATGTCATCTGCTCATTTG 2742
QY 1581 CATGCTGCGACGCGCTCATGTCACACCCCGTACTGACCCCGATGACCCCAAGATACAT 1640
Db 2743 GCTCAATATACCGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2802
QY 1641 TGTCTCCGTATACCCCAACATCAACGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1700
Db 2803 CATACCTGCGACACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2862
QY 1701 GCTCTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1760
Db 2863 GCTGCGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2922
QY 1761 CGAGGCGCAAGTTCAACACCTGACATGACCTTCTATTTCACTGCTGCTGCTGCTGCT 1820
Db 2923 TGAAGCCAAAGTTCAACACCTGACATGACCTTCTATTTCACTGCTGCTGCTGCTGCT 2982
QY 1821 CACCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1880
Db 2983 CCGCGCTGACGCGACACTTACGCAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3042
QY 1881 GCTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1940
Db 3043 GCGCGCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3102
QY 1941 CTACCGGAGGCGCAACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2000
Db 3103 CAAGCCTTCCCGGAACACATGAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3162
QY 2001 GAGGAGAC 2007
Db 3163 GGTGGCC 3169

```

```

RESULT 7
US-08-484-565-1
: Sequence 1, Application US/08484565
: Patent No. 5763569
: GENERAL INFORMATION:
: APPLICANT: Edward M. Brown
: APPLICANT: Steven C. Hebert
: APPLICANT: James E. Garrett, Jr.
: TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
: TITLE OF INVENTION: MOLECULES
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon

```

```

STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSO
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,565
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEPHONE: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5275 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 515..3769
OTHER INFORMATION:
US-08-484-565-1
Query Match 8.6%; Score 172.6; DB 1; Length 5275;
Best Local Similarity 50.2%; Pred. No. 3.4e-28;
Matches 455; Conservative 0; Mismatches 449; Indels 3; Gaps 1;

```

QY 1281 TCAGTCGAGTGGGGGCCCCATGCTCTTCATGCTGACACTGCTGCTGCTGATACAT 1340
 Db 2446 CAAGGCCACCAACGGGAGCTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTC 2505
 QY 1341 GGTGTCCTCCGCTGACGTGGGGGCCCCAGAGTCTCCACTGCTCTGCGCCGACGGCCT 1400
 Db 2506 CAGCTCCCTGCTC 2565
 QY 1401 CTTCCTCCCTGCTC 1460
 Db 2566 CTTCGCTGATCAGCTC 2625
 QY 1461 CTGCGCTTCAAGATGGCCACGCTTCCACGCGCTTACAGTACTGCTGCTGCTGCTACA 1520
 Db 2626 CTTGCTGCTTGA---GGCCAAAGATTCACACAGCTTCCACCGGAAGTGGGGGCTCA 2682
 QY 1521 GGGGCCCCAGTCTCTATGAGATTTATCAGGCTACTCAAAATGCTCAATGTGTAATTGG 1580
 Db 2683 CTTGCTGCT 2742
 QY 1581 CATGCTGGCAAGGCTCTAGTCCACCCCGTACTGACCCGATGACCCCAAGATCACAAAT 1640
 Db 2743 GCTCAATACAGCGCCGCCCTGAGCTACCGCAACGAGCTGAGAGACGATCATCTT 2802
 QY 1641 TGTCTCTGTACCCCACTACGCAACAGCTGCTGTTCAACACGAGCTGAGCTGCT 1700
 Db 2803 CATCACTGCGCACAGGAGCTGCTCATGCGGCTGGGCTTCTCTGATGAGCTACACTGCTT 2862
 QY 1701 GCTCTGATGCTGGGCTTTCAGCTCTGCTCTACATGAGGCAAGAGCTGCCCACTACAA 1760
 Db 2863 GCTGCGCGCCATCTGCT 2922
 QY 1761 CGAGGCCAAGTTCACACCTCAGCATGACTTCTATTTCATCTCATCTCCCTCTG 1820
 Db 2923 TGAAGCCAGTTCATCACTCTGACATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2982
 QY 1821 CACCTTATGCT 1880
 Db 2983 CCCCCCTACGCCAGCACTTACGGCAAGTTCGCTCTGCGCTGAGGATGAGCTACCT 3042
 QY 1881 GCTCAACTCTCTGCGCCATCAGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1940
 Db 3043 GCGCGCCAGCTTGTGCT 3102
 QY 1941 CTACCCGAGGCAACAGCGCCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2000
 Db 3103 CAAGCTTCCCGGAACACATCAGAGAGTCTGCTGCAACCGCGGACACGCTTCA 3162
 QY 2001 GAGGAC 2007
 Db 3163 GGTGGCC 3169

RESULT 8
 US-08-480-751-1
 ; Sequence 1, Application US/08480751
 ; Patent No. 5858684

GENERAL INFORMATION:

; APPLICANT: Edward F. Nemeth
 ; APPLICANT: Edward M. Brown
 ; APPLICANT: Steven C. Hebert
 ; APPLICANT: Forrest H. Fuller
 ; APPLICANT: James E. Garrett, Jr.
 ; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 ; TITLE OF INVENTION: MOLECULES
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lyon & Lyon
 ; STREET: First Interstate World Center
 ; STREET: Suite 4700
 ; STREET: 633 West Fifth Street
 ; CITY: Los Angeles

; STATE: California
 ; COUNTRY: USA
 ; ZIP: 90071
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" diskette, 1.44 Mb storage
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: FASTSEQ
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/480,751
 ; FILING DATE: 7 June 1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; PRIOR APPLICATION DATA: including application
 ; PRIOR APPLICATION DATA: described below: 9
 ; APPLICATION NUMBER: 08/353,784
 ; FILING DATE: 9 December, 1994
 ; APPLICATION NUMBER: PCT/US/94/12117
 ; FILING DATE: 21 October, 1994
 ; APPLICATION NUMBER: U.S. 08/292,827
 ; FILING DATE: 23 August, 1994
 ; APPLICATION NUMBER: U.S. 08/141,248
 ; FILING DATE: 22 October, 1993
 ; APPLICATION NUMBER: U.S. 08/009,389
 ; FILING DATE: 23 February, 1993
 ; APPLICATION NUMBER: U.S. 08/017,127
 ; FILING DATE: 12 February, 1993
 ; APPLICATION NUMBER: U.S. 07/934,161
 ; FILING DATE: 21 August, 1992
 ; APPLICATION NUMBER: U.S. 07/834,044
 ; FILING DATE: 11 February, 1992
 ; APPLICATION NUMBER: U.S. 07/749,451
 ; FILING DATE: 23 August, 1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Hebert, Sheldon O.
 ; REGISTRATION NUMBER: 38,179
 ; REFERENCE/DOCKET NUMBER: 213/004
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (213) 489-1600
 ; TELEFAX: (213) 955-0440
 ; TELEX: 67-3510
 ; INFORMATION FOR SEO ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 5275 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA to mRNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 515..3769
 ; OTHER INFORMATION:
 ; US-08-480-751-1

Query Match 8.6%; Score 172.6; DB 2; Length 5275;
 Best Local Similarity 50.2%; Pred. No. 3.4e-28;
 Matches 455; Conservative 0; Mismatches 449; Indels 3; Gaps 1;

QY 1101 TGAATGCCCAATATACAGAGTGTCTCTACAGAGAGTACCTCTCTCTCTCTCTCTCTCT 1160
 Db 2266 TAAGTGCCCTGATGACTTGTGTGCTCAATGAGAACACACTTCTGCAATGCCAAGGAGAT 2325
 QY 1161 GGTCTTCCTGGAATGAGATGAGGACACCATGCTGTGCGCTGCTGCTGCTGCTGCTG 1220
 Db 2326 CGAGTTCTGTCTGTGACGACGACCTTGGGATGACACTACAGCTCTTCTGCTGCTGG 2385
 QY 1221 CTTCCTGAGCACTGCGCATCTCTGTGATATTCTGAGGCACTTCCAGACACCATAGT 1280
 Db 2386 CATTTCTCACAAGCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2445
 QY 1281 TCAGTCGAGTGGGGGCCCCATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1340
 Db 2446 CAAGGCCACCAACGGGAGCTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2505

```

QY 1341 GGTGTCCTCCGCTGACGTGGGGCCCCCAAGCTCTCCACCTGCTCTGCCCCAGGCCCT 1400
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2506 CAGCTCCCTGTTCTTCATCGGGAGCCCCAGGACTGGAGCTGCTCCGCGCACCCGCGC 2565
QY 1401 CTTTCCCTCTCTCCACAAATTGATCTCGTATTCGCGGGCTTCCTTTCCAGATGCT 1460
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2566 CTTTGGCATCGCTTGTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2625
QY 1461 CTGCGCTTCAAGATGGCCGCGCTTCCACGCGCTTACAGCTTACGCTGCTGCTGCTGCT 1520
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2626 CTTGCTGTTTGA---GGCAAGATTCCACAGCTTCCACCGGAGAGTGTGGGGCTCAA 2682
QY 1521 GGGGGCTTACGCTCTATGAGCTTATACGCTTACTCAAAATGCTGCTGCTGCTGCTGCT 1580
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2683 CTTGAGTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2742
QY 1581 CATGCTGACAGGCTTCACTGCTTCCACCGCTTACCTGCTGCTGCTGCTGCTGCTGCTGCT 1640
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2743 GCTCAATACAGGCGCCCTCGAGCTACCGCAACGAGCTGGAGAGAGATCATCTT 2802
QY 1641 TGTCTCTGTAACTCCCACTACCGAGAGCTGCTGCTTCAACACAGCTGAGCTGCTGCT 1700
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2803 CATCACTGACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2862
QY 1701 GCTCAGTGTGGGTTTACGCTTACGCTTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1760
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2863 GCTGCGCCCATCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2922
QY 1761 CGAGGCAAGTTCATCAGCTTACGCTTACGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTG 1820
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2923 TGAAGCCAAAGTTCATCAGCTTACGCTTACGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2982
QY 1821 CACCTTCATGCTGCTTACGAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1880
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2983 CCCCCCTACGCGACACTTACGAGAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3042
QY 1881 GCTCAACCTCTGCGCATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1940
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3043 GCGCGCCAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3102
QY 1941 CTACCCGAGGCGCAACAGCGCCGCTTACTTCAACAGCATGATCCAGGCTTACACCATGAG 2000
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3103 CAAGCTTCCGAGACACATGAGAGAGTGGCTGACAGCAGCGGGGACACGCTTCAA 3162
QY 2001 GAGGAC 2007
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 3163 GGTGGCC 3169

```

```

RESULT 9
US-08-943-986-1
Sequence 1, Application US/08943986
Patent No. 5962314
GENERAL INFORMATION:
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESS: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible

```

```

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943,986
FILING DATE: 03-OCT-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/484,565
FILING DATE: 7-June-1995
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO. 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5275 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 515..3769
OTHER INFORMATION:
US-08-943-986-1

Query Match      8.6%  Score 172.6; DB 2; Length 5275;
Best Local Similarity 50.2%; Pred. No. 3.4e-28;
Matches 455; Conservative 0; Mismatches 449; Indels 3; Gaps 1;

QY 1101 TGAATGCCGGAATTAACGAGTGTGCTTACAGAGTGAAGTCACTTCTGCTTCAAGCGGAGCT 1160
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2266 TAAAGCCCTGATGACTCTGTGCTCAATGAGAACCAACTTCTGATGCGCAAGAGAT 2325
QY 1161 GGTCTTCTGGAATGAGATGAGGACACCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1220
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2326 CAGATTTCTGTGCTGAGCCAGCCCTTGGGAGCTGACATCAAGCTTGTGCTGCTGCTGCTG 2385
QY 1221 CTTTCCAGACCCCTGGCCATCCTGCTGATATTTGAGAGCACTTCCAGACCCCATAGT 1280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2386 CATTTTCTGACAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2445
QY 1281 TCGTTCGCTGGGGGCCCATGTGCTTCTGATGCTGACATGCTGCTGCTGCTGCTGCTGCTGCT 1340
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2446 CAAGGCCAACACCGGAGCTCTCTATCTCTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2505
QY 1341 GGTGTCCTCCGCTGATCGGGGCGCCCAAGTCTTCCAGCTTCTGCTGCTGCTGCTGCTGCTGCT 1400
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2506 CAGCTCCCTGTTCTTATCGGGGAGCCCGAGGAGTGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 2565
QY 1401 CTTTCCCTCTGCTTCAAAATTGATCTCTGATTCGCGCTGCTGCTTCTTCCAGATGCT 1460

```


OY 1521 GGGGCGCTACGTCCTATGATGATATGACGGTACTCAAAATGTCATGTTGTAATTGG 1580
 DB 2683 CTTGAGTTCCTGCTGCTGCTTCTCTGACCTTATGCGATTTGATGTCATTTGG 2742
 OY 1581 CATGCTGACGAGGCTCTGATGCTCCACCCCGTACTGACCCCGATGACCCCAAGATCAAT 1640
 DB 2743 GCTCAATACAGGCGCCCTGAGCTACCGCAACGAGCTGGAGGAGATGATCATCTT 2802
 OY 1641 TGTCTCTGTATACCCCACTACCGCAAGCCTGCTGTTCAACACCCAGCTGGACCTGCT 1700
 DB 2803 CATCACTGCTCCACGAGGCTGCTCATGCGCTGGGCTTCCGATGCGCTACACCTGCTT 2862
 OY 1701 GCTCAAGTGTGGTGTTCAGCTGCGCTACATGGGCAAGAGCTGGCCACCACTACAA 1760
 DB 2863 GCTGGCGCCGATCTGCTTCTTCTTCTGCTTCAAGTCCGGAAGCTGGCAGAGACTTCA 2922
 OY 1761 CGAGGCGCAAGTTCATACCCCTGACGATGACCTTCTATTTACCTCATCCGCTCCCTG 1820
 DB 2923 TGAAGCCAAAGTTCATACCTTACGATGCTCATCTTCTTCAATCGTGTGATCTCTTCA 2982
 OY 1821 CACCTTCATGTCCTGCTACAGGGGGTCTGCTGCTACCATCGTGACCTTGTGCTACGT 1880
 DB 2983 CCGCGCTACGCGACGACTTACGCAAGTCTGCTGCTGCGGAGGATGCGCATCCT 3042
 OY 1881 GCTCAACCTCTGCGCATACGCTGGGCTACTTTCGCGCCCAAGTCTACATGCTCTT 1940
 DB 3043 GCGCGCCAGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3102
 OY 1941 CTACCGGAGGCGCAACAGCCCGCTACTTCAACAGCATGATGACAGGCTGACCATGAG 2000
 DB 3103 CAAGGCTTCCGGAACACCATGAGAGGAGTGGCGTGCAGACCGCGGACACGCTTCAA 3162
 OY 2001 GAGGAGC 2007
 DB 3163 GGTGGCC 3169

RESULT 11
 US-08-484-719B-1
 ; Sequence 1, Application US/08484719B

; Patent No. 6031003
 ; GENERAL INFORMATION:
 ; APPLICANT: Edward F. Nemeth, Edward M.
 ; APPLICANT: Brown, Steven C. Hebert,
 ; APPLICANT: Bradford C. Van Wagonen,
 ; APPLICANT: Manuel F. Balandrin,
 ; APPLICANT: Forrest H. Fuller, Eric G.
 ; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 ; MOLECULES
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lyon & Lyon
 ; STREET: First Interstate World Center
 ; STREET: Suite 4700
 ; STREET: 633 West Fifth Street
 ; CITY: Los Angeles
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 90071
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: MS Word
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/484,719B
 ; FILING DATE: 7 June, 1995
 ; CLASSIFICATION: 514
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/353,784
 ; FILING DATE: 9 December, 1994

APPLICATION NUMBER: PCT/US/94/12117
 FILING DATE: 21 October, 1994
 APPLICATION NUMBER: U.S. 08/292,827
 FILING DATE: 23 August, 1994
 APPLICATION NUMBER: U.S. 08/141,248
 FILING DATE: 22 October, 1993
 APPLICATION NUMBER: U.S. 08/009,389
 FILING DATE: 23 February, 1993
 APPLICATION NUMBER: U.S. 08/017,127
 FILING DATE: 12 February, 1993
 APPLICATION NUMBER: U.S. 07/934,161
 FILING DATE: 21 August, 1992
 APPLICATION NUMBER: U.S. 07/834,044
 FILING DATE: 11 February, 1992
 APPLICATION NUMBER: U.S. 07/749,451
 FILING DATE: 23 August, 1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Douglas C. Murdock
 REGISTRATION NUMBER: 37,549
 REFERENCE/DOCKET NUMBER: 213/007
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5275 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 515..3769
 US-08-484-719B-1

Query Match 8.6%; Score 172.6; DB 3; Length 5275;
 Best Local Similarity 50.2%; Pred. No. 3.4e-28;
 Matches 455; Conservative 0; Mismatches 449; Indels 3; Gaps 1;

OY 1101 TGAATCCCGAATTAAGATGCTGCTTACCAAGATGACCTCTGCTTCAAGCGCAGCT 1160
 DB 2266 TAAAGTCCCTGATGACTTCTGCTCAATGAGAACCAACTTCTGCTCAACGCAAGAGAT 2325
 OY 1161 GGTCTTCTGGAATGATGAGGACCAACCATCTGCTGCGCCCTGCTGCGCCCTGGG 1220
 DB 2326 CAGATTCTGCTGACCGACCCCTTGGGATGCGACTACAGCTCTTGTGCTGCTGGG 2385
 OY 1221 CTTCCTACGACCTTGGCCATCTGCTGATATTTGAGAGCACTTCCAGACACCATAGT 1280
 DB 2386 CATTTTCTCAAGGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2445
 OY 1281 TCGCTGCGCTGGGCGCCCATGCTTCTGATGCTGACACTGCTGCTGCTGCTGCTGCT 1340
 DB 2446 CAAGGCCACCAACCGGAGCTCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2505
 OY 1341 GGTGCTCCCGGTACGAGGCGCCCGCAAGTCTTCAACTGCTGCTGCTGCTGCTGCT 1400
 DB 2506 CAGTCCCTGCTTCTATCGGGGAGCCCGGAGCTGAGAGTGCCTGCGCGCACCGGC 2565
 OY 1401 CTTCCTGCTGCTTCAATATGATGCTCTGATGCGCGCTGCTTCTTCTTCTTCTTCT 1460
 DB 2566 CTTTGGCATGACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2625
 OY 1461 CTGCGCTTCAAGATGCGCGCTTCCACGCGCTTACAGCTTACAGCTTGGCTGCTTACCA 1520
 DB 2626 CCGTGTGTTGA---GCCAAGATTCACACAGCTTCCACCGGAGTGTGGGGGCTCAA 2682
 OY 1521 GGGGCGCTACGCTCTATGAGCTTATACAGGTACTCAAAAAGTCATTTGTAATTGG 1580
 DB 2683 CCGAGATTCGCTGCTGCTTCTTCTGACCTTCAATGAGAGATGATGCTGCTGCTTGG 2742
 OY 1581 CATCTGACAGGCGCTGATGCCACCCCGTACTGACCCGATGACCCCAAGATCAAT 1640

Db 2743 GCTCAATACAGCGCCCTTCAGCTACCGAACACAGAGCTGAGAGACGATCATCTT 2802
 QY 1641 TGTCTCTGTAAACCCCAACACACACAGCTGCTTTCACACACAGCTGAGACCTGCT 1700
 Db 2803 CATACCTGCGACAGGGGCTGCTGCTGAGGGGCTGCTGCTGCTGCTGCTGCTGCTGCT 2862
 QY 1701 GCTCTCAGTGTGGGTTCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1760
 Db 2863 GCTGGGGCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2922
 QY 1761 CGAGCCCAAGTTCATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1820
 Db 2923 TGAAGCCCAAGTTCATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2982
 QY 1821 CACCTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1880
 Db 2983 CCGCGCTTACGCGACGCTTACGCGAGTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3042
 QY 1881 GCTCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1940
 Db 3043 GCGCGGACGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3102
 QY 1941 CTACCCGAGGCGACACAGCGCCGCTTCTTACAGACATGATTCAGGCTTACACCATGAG 2000
 Db 3103 CAAGCCTTCCCGAACAACATCGAGAGGTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3162
 QY 2001 GAGGGAC 2007
 Db 3163 GGTGGCC 3169

RESULT 12

US-08-484-159-1

Sequence 1, Application US/08484159

Patent No. 6313146

GENERAL INFORMATION:

APPLICANT: Bradford C. Van Wagenen

APPLICANT: Manuel F. Balandrin

APPLICANT: Eric G. Del Mar

APPLICANT: Edward F. Nemeth

TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESSES:

ADDRESSEE: Lyon & Lyon

STREET: First Interstate World Center

STREET: Suite 4700

CITY: Los Angeles

STATE: California

COUNTRY: USA

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: FASTSEQ

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,159

FILING DATE: 7 June, 1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA: including application

PRIOR APPLICATION DATA: described below: 9

APPLICATION NUMBER: 08/353,784

FILING DATE: 9 December, 1994

APPLICATION NUMBER: PCT/US/94/12117

FILING DATE: 21 October, 1994

APPLICATION NUMBER: U.S. 08/292,827

FILING DATE: 23 August, 1994

APPLICATION NUMBER: U.S. 08/141,248

FILING DATE: 22 October, 1993

APPLICATION NUMBER: U.S. 08/009,389
 FILING DATE: 23 February, 1993
 APPLICATION NUMBER: U.S. 08/017,127
 FILING DATE: 12 February, 1993
 APPLICATION NUMBER: U.S. 07/934,161
 FILING DATE: 21 August, 1992
 APPLICATION NUMBER: U.S. 07/834,044
 FILING DATE: 11 February, 1992
 APPLICATION NUMBER: U.S. 07/749,451
 FILING DATE: 23 August, 1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Heber, Sheldon O.
 REGISTRATION NUMBER: 38,179
 REFERENCE/DOCKET NUMBER: 214/101
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELETYPE: 67-3510
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5275 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 MOLECULE TYPE: linear
 TOPOLOGY: linear
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 515..3769
 OTHER INFORMATION:
 US-08-484-159-1

Query Match

8.6%; Score 172.6; DB 4; Length 5275;

Best Local Similarity 50.2%; Pred. No. 3.4e-28;

Matches 455; Conservative 0; Mismatches 449; Indels 3; Gaps 1;

QY 1101 TGAATGCCGGAATPACAGATGCTTCTTACAGAGTGAAGTCTTCTTCTTCAAGCGCATGCT 1160
 Db 2266 TAACTGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2325
 QY 1161 GGTCTTCCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1220
 Db 2326 CGAGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2385
 QY 1221 CTTCCTCAGACACCTGCGCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1280
 Db 2386 CATTTCTCAGACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2445
 QY 1281 TCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCTGCGCT 1340
 Db 2446 CAAGGCCACCAACCGGAGGCT 2505
 QY 1341 GGTGTCGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1400
 Db 2506 CAGTCTCCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2565
 QY 1401 CTTTCCTCCTGCTGCTTCAATTTGATCTCTCTGATGCGGCTGCTTCTTCTTCTTCTTCTTCTTCT 1460
 Db 2566 CTTTGGATCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2625
 QY 1461 CTGCGCTTCAAGATGCGCAGCGCTTCCAGCGGCTTCAAGCTACGATGCTGCTGCTGCTGCTGCTGCT 1520
 Db 2626 CTGCTGCTTGA--GGCCAAAGATTCACACAGCTTCCACCGGAGGAGGCTGCTGCTGCTGCTGCTGCT 2682
 QY 1521 GGGGCGCTTACGCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1580
 Db 2683 CTGCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2742
 QY 1581 CAGCTGCGACGCGCTGATGCGACCGCGCTTACGCTTACGACCGCGATGACCGCAAGATCACAAT 1640
 Db 2743 GCTCAATACAGCGCGCGCTGCGATGCTGCGATGCTGCGATGCTGCGATGCTGCGATGCTGCTGCTGCT 2802
 QY 1641 TGTCTCTGTAAACCCCAACACACAGCTGCTTTCACACACAGCTGCTGCTGCTGCTGCTGCTGCTGCT 1700

Db 2803 CATCAGCTGCCAGGAGGCTGCTCATGCGCTGGGCTTCTGATCGGCTACACCTGCT 2862
QY 1701 GCTCTAGTGTGGTTTACCTTCGCTTACATGGGCAAAAGCTGCCACCACTACAA 1760
Db 2863 GCTGGCGCCAGTCTGCTTCTTCCGCTTCAAGTCCGGAAGTGGCAGAACTTCAA 2922
QY 1761 CGAGGCGAGTTTCACTACAGCTTACATGATCTTATTTACCTTATCCGCTCCCTG 1820
Db 2923 TGAAGCCAGTTTCACTACAGCTTACATGATCTTATTTACCTTATCCGCTCCCTG 2982
QY 1821 CACCTTATGCTGCTTACAGGCGGCTGCTGATCAGCAATCCTGATCTTGTACATG 1880
Db 2983 CCGGCTTACGCTGCTTACAGGCGGCTGCTGATCAGCAATCCTGATCTTGTACATG 3042
QY 1881 GCTAACCTTCTGCTTACAGGCGGCTGCTGATCAGCAATCCTGATCTTGTACATG 1940
Db 3043 GCGGCGCTTCTGCTTACAGGCGGCTGCTGATCAGCAATCCTGATCTTGTACATG 3102
QY 1941 CTACCCGAGCGCAAGCGGCTGCTGATCAGCAATCCTGATCTTGTACATG 2000
Db 3103 CAGGCTTCCGGAACACATCAGAGGCTGCTGATCAGCAATCCTGATCTTGTACATG 3162
QY 2001 GAGGAC 2007
Db 3163 GGTGGC 3169

RESULT 13

US-08-485-588-3
Sequence 3, Application US/08485588

Patent No. 5688938
GENERAL INFORMATION:

APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: Forrest H. Fuller
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIDIUM RECEPTOR-ACTIVE
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: FASTSEQ

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/485,588

FILING DATE: 7 June, 1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA: including application

PRIOR APPLICATION DATA: described below: 9

APPLICATION NUMBER: 08/353,784

FILING DATE: 9 December, 1994

APPLICATION NUMBER: PCT/US/94/12117

FILING DATE: 21 October, 1994

APPLICATION NUMBER: U.S. 08/292,827

FILING DATE: 23 August, 1994

APPLICATION NUMBER: U.S. 08/141,248

FILING DATE: 22 October, 1993

APPLICATION NUMBER: U.S. 08/009,389

FILING DATE: 23 February, 1993

APPLICATION NUMBER: U.S. 08/017,127

FILING DATE: 12 February, 1993

APPLICATION NUMBER: U.S. 07/934,161

FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Hebert, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3809 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 373..3606
OTHER INFORMATION:
US-08-485-588-3

Query Match 8.3%; Score 167; DB 1; Length 3809;

Best Local Similarity 49.8%; Pred. No. 5e-27;

Matches 450; Conservative 0; Mismatches 450; Indels 3; Gaps 1;

QY 1103 AATGCCGAATACGAGAGTGTCTTACAGAGTGAAGCTTCTTCAAGGCGCAGTGG 1162
Db 2123 AGTCCCGAGATGATCTTGTGTCATGAGAAACACACCTCTGCTTCCAGAGAGATCG 2182
QY 1163 TCTTCTGGAATGGCATGAGGACCAACCATGCTGCTGCTGCTGCTGCTGCTGCT 1222
Db 2183 AGTTCTGCTGAGACGAGGACCTTGGGATGCGACTACACCTCTTCTGCTGCTGCT 2242
QY 1223 TCTTCTGAGACCTTGGCATCTGCTGATGATGCTGAGGACCTTCCAGACCCATAGTTC 1282
Db 2243 TTTTCTGAGACGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2302
QY 1283 GCTGCGTGGGGGCGCCCATGCTTCTGATGCTGACACTGCTGCTGCTGCTGCTGCT 1342
Db 2303 AGGCCACCAACCGAGAGAGTCTTCTTACCTTCTTCTTCTTCTTCTTCTTCTTCT 2362
QY 1343 TGTCTCCGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1402
Db 2363 GCTCCCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2422
QY 1403 TTCCCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1462
Db 2423 TTGGCATGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2482
QY 1463 GCGCTTCAAGATGCGCAAGCGCTTCCAGCGCTTACAGTACGCTTCCGCTACAGG 1522
Db 2483 TGTGTTTGA---GGCCAAAGTCCCAACAGCTTCCAGCGAGTGGGGGCGTCAAC 2539
QY 1523 GCGCTTACGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 1582
Db 2540 TGCAGTCTCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2599
QY 1583 TGTGCGCAGCGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1642
Db 2600 TCTTACACCGCGCCCGCTTCAAGCTTACGCAACAGAGCTGAGAGATGATCATCTTCA 2659
QY 1643 TCTCTTGAACCCCAATACGCAAGAGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1702
Db 2660 TCAAGTGCAGAGAGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2719
QY 1703 TCTTCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1762
Db 2720 TGGCTGCATCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 2779

Db 2840 CAGCCTATGACAGACCTATGAGGAGTTGTCTGCGGTAGAGGTGATGTCATCCG 2899
 Oy 1883 TCACCTCTGCGCATGACCTGCGCTACTTGCGCCCAAGTGTACATGATCTCTTCT 1942
 Db 2900 CAGCAGCTTGGCTGTGCTGGCGGTCTCTTCTCAACAGATCTACATCTCTTCA 2959
 Oy 1943 ACCCGAGGCAACAGCGCGCTTCAACAGCATGATCCAGGCTACACCATGAGA 2002
 Db 2960 AGCCATCCGCAACACATGAGAGGTGCTTGCAGACCGCAGCTCAGCTTCAAGG 3019
 Oy 2003 GGG 2005
 Db 3020 TGG 3022
 RESULT 15
 US-08-480-751-3
 ; Sequence 3, Application US/08480751
 ; Patent No. 5858684
 ; GENERAL INFORMATION:
 ; APPLICANT: Edward F. Nemeth
 ; APPLICANT: Edward M. Brown
 ; APPLICANT: Steven C. Hebert
 ; APPLICANT: Forrest H. Fuller
 ; APPLICANT: James E. Garrett, Jr.
 ; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 ; TITLE OF INVENTION: MOLECULES
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Lyon & Lyon
 ; STREET: First Interstate World Center
 ; STREET: Suite 4700
 ; CITY: Los Angeles
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 90071
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: FASTSEQ
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/480,751
 ; FILING DATE: 7 June, 1995
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; PRIOR APPLICATION DATA: including application
 ; PRIOR APPLICATION DATA: described below: 9
 ; APPLICATION NUMBER: 08/353,784
 ; FILING DATE: 9 December, 1994
 ; APPLICATION NUMBER: PCT/US/84/12117
 ; FILING DATE: 21 October, 1994
 ; APPLICATION NUMBER: U.S. 08/292,827
 ; FILING DATE: 23 August, 1994
 ; APPLICATION NUMBER: U.S. 08/141,248
 ; FILING DATE: 22 October, 1993
 ; APPLICATION NUMBER: U.S. 08/009,389
 ; FILING DATE: 23 February, 1993
 ; APPLICATION NUMBER: U.S. 08/017,127
 ; FILING DATE: 12 February, 1993
 ; APPLICATION NUMBER: U.S. 07/934,161
 ; FILING DATE: 21 August, 1992
 ; APPLICATION NUMBER: U.S. 07/834,044
 ; FILING DATE: 11 February, 1992
 ; APPLICATION NUMBER: U.S. 07/749,451
 ; FILING DATE: 23 August, 1991
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Heber, Sheldon O.
 ; REGISTRATION NUMBER: 38,179
 ; REFERENCE/DOCKET NUMBER: 213/004
 ; TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3809 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA to mRNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 373..3606
 ; OTHER INFORMATION:
 US-08-480-751-3
 Query Match 8.3%, Score 167; DB 2; Length 3809;
 Best Local Similarity 49.8%; Pred. No. 5e-27;
 Matches 450; Conservative 0; Mismatches 450; Indels 3; Gaps 1;

Oy 1103 AATGCCGATATACAGAGTGTCTTACAGAGAGACCTCTGCTTCAAGGGGAGCTGG 1162
 Db 2123 AGTGCCAGATGACTTCTGCTCAATGAGAACACCTCTGCTTCAAGAGATCG 2182
 Oy 1163 TCTTCGTAATGATGATGAGGACACCATGCTGTGGCCCTGCTGGCGGCTGGGCT 1222
 Db 2183 AGTTCTGTGCGAGGAGGCGCTTGGGATCGACTACCTCTTGGCGCTGGGGA 2242
 Oy 1223 TCTTCAGACCTGAGCATCTGATGATATCTGAGGACACTTCCAGACACCATAGTTC 1282
 Db 2243 TTTTCTGACAGCCTTGTGGGGGTGTGTTATCAAGTTCGCAACACCATGTCGA 2302
 Oy 1283 GCTGCGTGGGGGCCCATGTGCTTCTGATGCTACACTGTGCTGTGGCATCATGG 1342
 Db 2303 AGGCCACCAACCGAGAGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 2362
 Oy 1343 TGTGCGGATGATGATGAGGCGGCCCAAGGATCTCCAGCTTGCAGGCGGCTCT 1402
 Db 2363 GCTCCCTGTCTTCAATCGGGGAGGCCCGCAGACTGAGACCTGCGCCGCGCGCT 2422
 Oy 1403 TTCCCTGTGCTTCAAAATTTGATCTCTGATGCGCGGTCTTTCAGATGCTCT 1462
 Db 2423 TTGGCATGATCTCTGCTGCTGATCTATGATGATGATGATGATGATGATGATG 2482
 Oy 1463 GCGCCTTCAAGATGAGCGCGCTTCCAGCGGCTTCAAGCTACTGAGTCCGCTAC 1522
 Db 2483 TGTGTTTGA---GGCCAGATCCCGACAGCTTCCACCGCAAGTGTGGGGCTCAAC 2539
 Oy 1523 GGCCTACGATCTATGATGATTTATCAGGATCTCAAAATGCTATGTTGTAATGGCA 1582
 Db 2540 TGCAATTCCTGTGTTTTCCTGCTGACCTTCAATGATGATGATGATGATGATG 2599
 Oy 1583 TGTGCGACGGCTTATGTCACACCCCGTACTGACCCCGATGAGCCCAAGATCAATTG 1642
 Db 2600 TCTACACCGCGCCCGCCCTCAAGCTACCGCAACAGAGCTGAGATGAGATCACTTCA 2659
 Oy 1643 TCTCTGTAACCCCAATACGCAACAGAGCTCTGTCAAGCAAGCTGAGCTGTGTC 1702
 Db 2660 TCACGTCCACAGAGGGCTCCCTCAATGAGCCCTGAGGCTCTGATGCGCTACACTTCTG 2719
 Oy 1703 TCTCAGTGTGAGGTTTCACTTGCCTTCAATGAGGCAAGAGTGGCCCACTACATCAAG 1762
 Db 2720 TGGCGCATCTGCTTCTTCTTGGCTTCAAGTCCCGGAGTGGCGGAGATCTCAATG 2779
 Oy 1763 AGCCAGATTAATCAACCTCAGATGACCTTCTATTTCACTCAATGCTCTCTCTGCA 1822
 Db 2780 AAGCCAGATTAATCACTTCAAGATGATCTTCTTCAATCTCTGATCTCTCTCAATTC 2839
 Oy 1823 CTTATATCTCTTCAAGCGGGGTGCTGATCATGATGATGATGATGATGATGATG 1882
 Db 2840 CAGCCTATGACAGACCTATGAGCAAGTTGTCTGCGGTATGATGATGATGATGATG 2899
 Oy 1883 TCACCTCTGCGCATGAGCTGTGGCTTCTGCGCCCAAGTGTATGATGATCTCTTCT 1942

Db 2900 CAGCCAGCTTTGGCTTGCTGCGCTGCACTCTTCAACAAGATCTACATCATCTCTTCA 2959
QY 1943 ACCCGAGCGCAACAGCGCGCTTACTTCAACAGCATGATCCAGGGCTACACCATGAGGA 2002
Db 2960 AGCCATCCCGCAACACCATGAGGAGGTGCGCTTGACAGCACCGCAGCTCACGCTTTCAAGG 3019
QY 2003 GGG 2005
Db 3020 TGG 3022

Search completed: May 23, 2003, 18:23:42
Job time : 156.947 secs

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 23, 2003, 12:46:24 ; Search time 248.116 Seconds
(Without alignments)
10697.149 Million cell updates/sec

Title: US-09-927-315-12

Perfect score: 2010

Sequence: 1 ataccctacagcgcctacag.....acacatgagggagactag 2010

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_NA:*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PC7_NEM_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEM_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US07_NEM_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEM_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PC7US_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEM_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEM_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEM_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEM_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	2010	100.0	2010	9	US-10-096-144-8
2	2010	100.0	2010	9	US-09-927-315-12
3	2010	100.0	2010	12	US-10-124-598-8
4	1928.6	96.0	2520	9	US-09-897-427A-5
5	1928.6	96.0	2520	9	US-10-035-045-23
6	1232.8	61.3	2993	9	US-10-036-144-3
7	1232.8	61.3	2993	9	US-09-927-315-10
8	1232.8	61.3	2993	12	US-10-124-598-3
9	1212.6	60.3	2532	9	US-10-096-144-4
10	1212.6	60.3	2532	9	US-09-927-315-11
11	1212.6	60.3	2532	12	US-10-124-598-4
12	392.8	19.5	2771	9	US-08-361-652-4
13	392.8	19.5	2771	9	US-09-927-315-4
14	386.4	19.2	2292	10	US-09-819-946-3
15	386.4	19.2	2526	10	US-09-819-946-3
16	386.4	19.2	2951	10	US-09-819-946-9
17	385.2	19.2	2526	9	US-09-897-427A-1
18	385.2	19.2	2526	9	US-09-799-629-16
19	385.2	19.2	2526	9	US-10-035-045-16

20	376.6	18.7	2333	9	US-09-361-652-6	Sequence 6, Appl1
21	376.6	18.7	2333	9	US-09-927-315-6	Sequence 6, Appl1
22	359.2	17.9	2579	9	US-09-361-652-5	Sequence 5, Appl1
23	359.2	17.9	2579	9	US-09-927-315-5	Sequence 5, Appl1
24	258	12.8	697	10	US-09-811-284-117	Sequence 117, App
25	218.6	10.9	2559	9	US-09-897-427A-5	Sequence 5, Appl1
26	218.6	10.9	2559	9	US-09-897-427A-7	Sequence 7, Appl1
27	218.6	10.9	2559	9	US-09-927-315-14	Sequence 14, Appl1
28	218.6	10.9	2559	9	US-10-035-045-3	Sequence 3, Appl1
29	218.6	10.9	3458	9	US-10-261-482-1	Sequence 1, Appl1
30	217	10.8	2559	10	US-10-282-837-15	Sequence 15, Appl1
31	217	10.8	2559	10	US-09-796-338A-15	Sequence 15, Appl1
32	217	10.8	3489	9	US-10-282-837-13	Sequence 13, Appl1
33	217	10.8	3489	10	US-09-796-338A-13	Sequence 13, Appl1
34	216.2	10.8	2553	9	US-09-799-629-3	Sequence 3, Appl1
35	190.4	9.5	8194	9	US-09-799-629-15	Sequence 15, Appl1
36	190.4	9.5	8194	9	US-10-035-045-15	Sequence 15, Appl1
37	187	9.3	4134	9	US-09-975-553-17	Sequence 17, Appl1
38	187	9.3	4134	9	US-10-125-792-1	Sequence 1, Appl1
39	187	9.3	4134	9	US-10-125-792-1	Sequence 1, Appl1
40	187	9.3	4134	9	US-10-270-795-17	Sequence 17, Appl1
41	184.2	9.2	1101	10	US-09-819-946-5	Sequence 5, Appl1
42	167	8.3	3361	12	US-10-002-854-1	Sequence 1, Appl1
43	159	7.9	3237	10	US-09-727-205-1	Sequence 7, Appl1
44	157	7.8	705	10	US-09-819-946-7	Sequence 1, Appl1
45	127.4	6.3	2577	9	US-09-799-629-13	Sequence 13, Appl1

ALIGNMENTS

RESULT 1
US-10-096-144-8
Sequence 8, Application US/10096144
Publication No. US20030022288A1
GENERAL INFORMATION:
APPLICANT: Zuker, Charles S.
APPLICANT: Adler, Jon Elliot
APPLICANT: Lindemeyer, Juergen
TITLE OF INVENTION: The Regents of the University of California
TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
FILE REFERENCE: 02307E-088720US
CURRENT APPLICATION NUMBER: US/10/096.144
CURRENT FILING DATE: 2002-03-08
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/361,631
PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-27
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/095,464
PRIOR FILING DATE: EARLIER FILING DATE: 1998-07-28
PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/112,747
NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8
LENGTH: 2010
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human G-protein coupled receptor (GPCR) B4
US-10-096-144-8
Query Match 100.0%; Score 2010; DB 9; Length 2010;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2010; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATCACTACAGCGCCATGAGCTGCGAGACAAGAGTCCGCTTGGCTG 60
DB 1 ATCACTACAGCGCCATGAGCTGCGAGACAAGAGTCCGCTTGGCTG 60
QY 61 CGTACCAACCCAGCGCCGACACACAGTCCGAGGCGGACGACTATGCTGCATTC 120
DB 61 CGTACCAACCCAGCGCCGACACACAGTCCGAGGCGGACGACTATGCTGCATTC 120

QY 121 CGCTGGAACTGATCATTTGCTGTGTGAGCAGCAGACCTATGGCCGCGACAATGGCCAG 180
 Db 121 CGCTGGAACTGATCATTTGCTGTGTGAGCAGCAGACCTATGGCCGCGACAATGGCCAG 180
 QY 181 CTGCTTGGCAGAGCGCGTGGCCGCGGACATCTGCATCGCTTTCAGAGAGCGTGGCC 240
 Db 181 CTGCTTGGCAGAGCGCGTGGCCGCGGACATCTGCATCGCTTTCAGAGAGCGTGGCC 240
 QY 241 ACACCTGAGCCCAACCAATGAGCTGAGAGAGCGCGCGGAGCGCGGAGCGCGGAGCGCGG 300
 Db 241 ACACCTGAGCCCAACCAATGAGCTGAGAGAGCGCGCGGAGCGCGGAGCGCGGAGCGCGG 300
 QY 301 GACAACTGAGCAGAGCAGCAGCAGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
 Db 301 GACAACTGAGCAGAGCAGCAGCAGCGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 360
 QY 361 TACCACTTCTTCAATGAGT 420
 Db 361 TACCACTTCTTCAATGAGT 420
 QY 421 GAGTCTGGGCGATCGACCGCGT 480
 Db 421 GAGTCTGGGCGATCGACCGCGT 480
 QY 481 TTCTGTGGGCGATCGACCGCGT 540
 Db 481 TTCTGTGGGCGATCGACCGCGT 540
 QY 541 GGGCCACAGCGTGGCG 600
 Db 541 GGGCCACAGCGTGGCG 600
 QY 601 GAGTGGCAACACTGCTGAGAGCG 660
 Db 601 GAGTGGCAACACTGCTGAGAGCG 660
 QY 661 GAGGT 720
 Db 661 GAGGT 720
 QY 721 CTCTGTGGGCGATCGACCGCGT 780
 Db 721 CTCTGTGGGCGATCGACCGCGT 780
 QY 781 GAGGAGATCTGGAAGTCACTTCACTCTCTGAGACCAACAACTTCTTTCGAGCCGCGAA 840
 Db 781 GAGGAGATCTGGAAGTCACTTCACTCTCTGAGACCAACAACTTCTTTCGAGCCGCGAA 840
 QY 841 GGGGAGCGT 900
 Db 841 GGGGAGCGT 900
 QY 901 TTCCAGAGCGT 960
 Db 901 TTCCAGAGCGT 960
 QY 961 CTGCAACCGT 1020
 Db 961 CTGCAACCGT 1020
 QY 1021 CAAAAGAGAGCGT 1080
 Db 1021 CAAAAGAGAGCGT 1080
 QY 1081 GGGACCTTCTTCAACCACTGATGT 1140
 Db 1081 GGGACCTTCTTCAACCACTGATGT 1140
 QY 1141 TTCTGTCTTCAAGGCGAGT 1200
 Db 1141 TTCTGTCTTCAAGGCGAGT 1200

QY 1201 GCCCTGTGGCGCGCGCTTCCCTCAGACCGCTGGGCATCTCTGATATTCGTGAGG 1260
 Db 1201 GCCCTGTGGCGCGCGCTTCCCTCAGACCGCTGGGCATCTCTGATATTCGTGAGG 1260
 QY 1261 CACTTCCAGACACCCATAGTTCGCTGGCTGGGGGCCCATATGTCTTCTGATGCTGACA 1320
 Db 1261 CACTTCCAGACACCCATAGTTCGCTGGCTGGGGGCCCATATGTCTTCTGATGCTGACA 1320
 QY 1321 CTGCTGTGGGCGATCATAGT 1380
 Db 1321 CTGCTGTGGGCGATCATAGT 1380
 QY 1381 TGGCTGTGGCGCAGGCGCTTCTTCCCTCAGACCGCTGGGCATCTCTGATATTCGTGAGG 1440
 Db 1381 TGGCTGTGGCGCAGGCGCTTCTTCCCTCAGACCGCTGGGCATCTCTGATATTCGTGAGG 1440
 QY 1441 GTGGCTTCTTCCAGATGCTGT 1500
 Db 1441 GTGGCTTCTTCCAGATGCTGT 1500
 QY 1501 AGCTACTGGGTCGCTACAGGGGCGCTTACGCTCTATGGCATTTATCAGGTAATCAGGTA 1560
 Db 1501 AGCTACTGGGTCGCTACAGGGGCGCTTACGCTCTATGGCATTTATCAGGTAATCAGGTA 1560
 QY 1561 ATGATCTGT 1620
 Db 1561 ATGATCTGT 1620
 QY 1621 GATGACCCCAAGATCACAATTTGTCTCTGTATACCCCACTACGCGCAACAGCGCTGTCT 1680
 Db 1621 GATGACCCCAAGATCACAATTTGTCTCTGTATACCCCACTACGCGCAACAGCGCTGTCT 1680
 QY 1681 AACACAGCGCTGGACGCTGT 1740
 Db 1681 AACACAGCGCTGGACGCTGT 1740
 QY 1741 GAGTGGCCACCACTACCAAGAGGCGCAAGTTATCACCCTCAGCATGACCTTCTATTTTC 1800
 Db 1741 GAGTGGCCACCACTACCAAGAGGCGCAAGTTATCACCCTCAGCATGACCTTCTATTTTC 1800
 QY 1801 AACTCATCCGCTTCCCTCTGT 1860
 Db 1801 AACTCATCCGCTTCCCTCTGT 1860
 QY 1861 GTGGACCTTGT 1920
 Db 1861 GTGGACCTTGT 1920
 QY 1921 AAGTCTACATGATCTCTTCTACCGGAGCGCAACAGCGCGCGCTTACTTCAACAGCATG 1980
 Db 1921 AAGTCTACATGATCTCTTCTACCGGAGCGCAACAGCGCGCGCTTACTTCAACAGCATG 1980
 QY 1981 ATCCAGGCGTACACCATGAGGAGGACTAG 2010
 Db 1981 ATCCAGGCGTACACCATGAGGAGGACTAG 2010

RESULT 2 US-09-927-315-12

; Sequence 12, Application US/09927315
 ; Publication No. US2003004045A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Zuker, Charles S.
 ; APPLICANT: Ryba, Nicholas J.P.
 ; APPLICANT: Nelson, Greg
 ; APPLICANT: Hoon, Mark A.
 ; APPLICANT: Chandrasekhar, Jayaram
 ; APPLICANT: Zhang, Yifeng
 ; APPLICANT: The Regents of the University of California
 ; APPLICANT: as represented by the Secretary of the
 ; Department of Health and Human Services
 ; TITLE OF INVENTION: Mammalian Sweet Taste Receptors

QY 1921 AAGTGTACATGATCCTTCTTACCCGAGGCAACAGGCCGCCCTACTTTCACAGCATG 1980
 Db 1921 AAGTGTACATGATCCTTCTTACCCGAGGCAACAGGCCGCCCTACTTTCACAGCATG 1980
 QY 1981 ATCCAGGGCTACACCATGAGGAGGACTAG 2010
 Db 1981 ATCCAGGGCTACACCATGAGGAGGACTAG 2010
 RESULT 3
 US-10-124-598-8
 ; Sequence 8, Application US/10124598
 ; Patent No. US20020119526A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Zuker, Charles S.
 ; APPLICANT: Adler, Jon Elliot
 ; APPLICANT: Lindemeyer, Juergen
 ; APPLICANT: The Regents of the University of California
 ; TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
 ; FILE REFERENCE: 02307E-088720US
 ; CURRENT APPLICATION NUMBER: US/10/124,598
 ; PRIOR FILING DATE: 2002-04-16
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/361,631
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1999-07-27
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/112,747
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 8
 ; LENGTH: 2010
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; OTHER INFORMATION: human G-protein coupled receptor (GPCR) B4
 ; OTHER INFORMATION: nucleotide sequence
 US-10-124-598-8
 Query Match 100.0%; Score 2010; DB 12; Length 2010;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2010; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 421 GAGTCTGTGGGCATTCAGACCCGGTCTTGACACAACTTCAGAGAGTGGGCACCTTGGGCACC 480
 QY 481 TTCTGGGCATTCACATTCAGAGAGGTGGCCATCCGGGGCTTCACTGATTCGGAGTGG 540
 Db 481 TTCTGGGCATTCACATTCAGAGAGGTGGCCATCCGGGGCTTCACTGATTCGGAGTGG 540
 QY 541 GGCCACAGGTGGGGCCGCCACCCCTCAGCAGACAGCCAGAGCTTACCTGCACACAG 600
 Db 541 GGCCACAGGTGGGGCCGCCACCCCTCAGCAGACAGCCAGAGCTTACCTGCACACAG 600
 QY 601 GAGTGCACAACTGCTGAAAGCCACTTGTCTTCAACACATTCAGAGCTCTGAGG 660
 Db 601 GAGTGCACAACTGCTGAAAGCCACTTGTCTTCAACACATTCAGAGCTCTGAGG 660
 QY 661 GAGGTGTCGCTACAGAGTGTACTGCGGGTATGCTGAGGCGCATTCGCTCAGAGC 720
 Db 661 GAGGTGTCGCTACAGAGTGTACTGCGGGTATGCTGAGGCGCATTCGCTCAGAGC 720
 QY 721 CTCTGGGCTGTGCAAAAGACCTGACCAAGAGGGTGTTCACCCCTGGCAGCTGCTT 780
 Db 721 CTCTGGGCTGTGCAAAAGACCTGACCAAGAGGGTGTTCACCCCTGGCAGCTGCTT 780
 QY 781 GAGGAGATCTGGAAGGTCACTTCACTCTCTGACCAACCAAACTCTTCGACCCGCA 840
 Db 781 GAGGAGATCTGGAAGGTCACTTCACTCTCTGACCAACCAAACTCTTCGACCCGCA 840
 QY 841 GGGAGCTGCTGCTGCACTTGAAGATTTGCAAGTGGCAATGGACCGAGCAGAAATCCC 900
 Db 841 GGGAGCTGCTGCTGCACTTGAAGATTTGCAAGTGGCAATGGACCGAGCAGAAATCCC 900
 QY 901 TTCCAGAGCGTGGCTCTTCACTTACCCCTGACAGCAGCTGAAGAAATCAGACATCT 960
 Db 901 TTCCAGAGCGTGGCTCTTCACTTACCCCTGACAGCAGCTGAAGAAATCAGACATCT 960
 QY 961 CTGACACACGTCACACAGATCCATATGTCATGTTCCAAAGAGTGGCAGTCCAGG 1020
 Db 961 CTGACACACGTCACACAGATCCATATGTCATGTTCCAAAGAGTGGCAGTCCAGG 1020
 QY 1021 CAAAAGAAAGAGCTGTGGGCATCCAGTGTGCTGCTGAGTGCATGATGCTTCCC 1080
 Db 1021 CAAAAGAAAGAGCTGTGGGCATCCAGTGTGCTGAGTGCATGATGCTTCCC 1080
 QY 1081 GGCACTTCTCTCAACCACTGAATGCGCCGAATPACAGTGTCTTCAAGAGTACACC 1140
 Db 1081 GGCACTTCTCTCAACCACTGAATGCGCCGAATPACAGTGTCTTCAAGAGTACACC 1140
 QY 1141 TTCGCTTCAAGGGGCGAGTGTGCTTCTGGAATGAGTGGACCCACCATGCGTGTG 1200
 Db 1141 TTCGCTTCAAGGGGCGAGTGTGCTTCTGGAATGAGTGGACCCACCATGCGTGTG 1200
 QY 1201 GCCCTGTGGCGCCCTGTGGCTTCTCAGACACCCCTGGCAATCTGTGATATTCTGAGG 1260
 Db 1201 GCCCTGTGGCGCCCTGTGGCTTCTCAGACACCCCTGGCAATCTGTGATATTCTGAGG 1260
 QY 1261 CACTTCCAGACACCATAGTTCGCTGGGCGGGGGCCCAATGCTTCCGATGCTGACA 1320
 Db 1261 CACTTCCAGACACCATAGTTCGCTGGGCGGGGGCCCAATGCTTCCGATGCTGACA 1320
 QY 1321 CTGCTGTGCTGATCATGATGATGCTCCGCTTACGTGGGGCGGCCAAGTCTCCACC 1380
 Db 1321 CTGCTGTGCTGATCATGATGATGCTCCGCTTACGTGGGGCGGCCAAGTCTCCACC 1380
 QY 1381 TGCCTGTGCGCCAGAGCCCTCTTCCCTCTGCTTCAAAATTTGACATCTCTGTATCGCC 1440
 Db 1381 TGCCTGTGCGCCAGAGCCCTCTTCCCTCTGCTTCAAAATTTGACATCTCTGTATCGCC 1440
 QY 1441 GTGCGTCTTCTTCAAGATGCTGCTGCGCTTCAAAATGAGCCAGCGCTTCCAGCGCTAC 1500
 Db 1441 GTGCGTCTTCTTCAAGATGCTGCTGCGCTTCAAAATGAGCCAGCGCTTCCAGCGCTAC 1500
 QY 1501 AGCTACTGGGTCCGCTACACAGGGGCGCTTACGCTCATATGSCATTTATACAGGTACTCAA 1560
 Db 1501 AGCTACTGGGTCCGCTACACAGGGGCGCTTACGCTCATATGSCATTTATACAGGTACTCAA 1560

QY 1561 ATGGCATGTGTAATGSGATGCTGGCAGGGCTCAGTCCCAACCCCTACTGACCC 1620
Db 1561 ATGGCATGTGTAATGSGATGCTGGCAGGGCTCAGTCCCAACCCCTACTGACCC 1620
QY 1621 GATGACCCCAAGATACAAATGTCCTGTAAACCCCACTACCCGACAGCCCTGCTGTC 1680
Db 1621 GATGACCCCAAGATACAAATGTCCTGTAAACCCCACTACCCGACAGCCCTGCTGTC 1680
QY 1661 AACACACACCTGGACCTGCTGCTCACTGAGTGGGTTTCAGCTTCCCTACATGGGCAAA 1740
Db 1661 AACACACACCTGGACCTGCTGCTCACTGAGTGGGTTTCAGCTTCCCTACATGGGCAAA 1740
QY 1741 GAGCTGCCCACTACAGAGAGGCAAGTTCACTACCTCAGACCTTCCTATTTTC 1800
Db 1741 GAGCTGCCCACTACAGAGAGGCAAGTTCACTACCTCAGACCTTCCTATTTTC 1800
QY 1801 ACCTCATCCGTCCTCCTGACCTTCACTGCTGCTACAGCGGGGCTGCTGTCACATC 1860
Db 1801 ACCTCATCCGTCCTCCTGACCTTCACTGCTGCTACAGCGGGGCTGCTGTCACATC 1860
QY 1861 GTGACCTCTGTGTCACCTGCTCAACCTCCTGGCCTACAGCTGGGCTACTGGCCCC 1920
Db 1861 GTGACCTCTGTGTCACCTGCTCAACCTCCTGGCCTACAGCTGGGCTACTGGCCCC 1920
QY 1921 AAGTGTACATGATCCTCTTACCCGGAGGCAACAGCCGCTACTTCAACAGCATG 1980
Db 1921 AAGTGTACATGATCCTCTTACCCGGAGGCAACAGCCGCTACTTCAACAGCATG 1980
QY 1981 ATCCAGGGCTACACCATGAGAGGAGACTAG 2010
Db 1981 ATCCAGGGCTACACCATGAGAGGAGACTAG 2010

RESULT 4

US-09-897-427A-3
; Sequence 3, Application US/09897427A
; Patent No. US20020160424A1
; GENERAL INFORMATION:
; APPLICANT: ADLER, JON ELLIOT
; APPLICANT: LI, XIADONG
; APPLICANT: STRAWSKI, LENA
; APPLICANT: XU, HONG
; APPLICANT: EHEVERRI, FERNANDO
; TITLE OF INVENTION: T1R HETERO-OLIGOMERIC TASTE RECEPTORS
; FILE REFERENCE: 078003-0282558
; CURRENT APPLICATION NUMBER: US/09/897,427A
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2520
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-897-427A-3

Query Match 96.0%; Score 1928.6; DB 9; Length 2520;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 9; Indels 27; Gaps 4;

QY 1 ATCACTACAGCGCCATCAGGATGAGCTGGAGACAAGTGCGCTTCCGGCTTTGCTG 60
Db 484 ATCACTACAGCGCCATCAGGATGAGCTGGAGACAAGTGCGCTTCCGGCTTTGCTG 543
QY 61 CGTACACACACCGCGCGACACACAGCTGAGGCCATGCTGACAGCTATGCTGACATTC 120
Db 544 CGTACACACACCGCGCGACACACAGCTGAGGCCATGCTGACAGCTATGCTGACATTC 603
QY 121 CGCTGAGACTGATATGCTGCTGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db 604 CGCTGAGACTGATATGCTGCTGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 663
QY 181 CTGCTTGGAGAGCGGTGGCGCGGCGAGACATGCTGATGCTTCCAGAGAGAGCGTGGCC 240

Db 664 CTGCTTGGAGAGCGGTGGCGCGGCGAGACATGCTGATGCTTCCAGAGAGAGCGTGGCC 723
QY 241 AACTGACACCCCAACAGACATGACGTCAAGAGAGCGCCAGCGGCTGATGACATTTG 300
Db 724 AACTGACACCCCAACAGACATGACGTCAAGAGAGCGCCAGCGGCTGATGACATTTG 783
QY 301 GACAAGCTGACAG 360
Db 784 GACAAGCTGACAG 843
QY 361 TACCACTTCTCAATGAGTGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 420
Db 844 TACCACTTCTCAATGAGTGTGCGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 903
QY 421 GAGTCTGGGCGATGACCGGCTGCTGACAACTCAGAGAGAGAGAGAGAGAGAGAG 480
Db 904 GAGTCTGGGCGATGACCGGCTGCTGACAACTCAGAGAGAGAGAGAGAGAGAGAG 963
QY 481 TTCTGGGCGATGACCGGCTGCTGACAACTCAGAGAGAGAGAGAGAGAGAGAGAG 540
Db 964 TTCTGGGCGATGACCGGCTGCTGACAACTCAGAGAGAGAGAGAGAGAGAGAGAG 1023
QY 541 GGGCCACAGGCTGGGCGCCGACCCCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
Db 1024 GGGCCACAGGCTGGGCGCCGACCCCTCAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1083
QY 601 GAGTGTGACAACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 660
Db 1084 GAGTGTGACAACTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1143
QY 661 GAGGCTGTGCTACAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 720
Db 1144 GAGGCTGTGCTACAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1203
QY 721 CTCTCGGCTGTGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
Db 1204 CTCTCGGCTGTGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1263
QY 781 GAGGAGATCTGGAAGGTAACTTCACTCTGAGACACCAAAATTTCTTGACACCCGAA 840
Db 1264 GAGGAGATCTGGAAGGTAACTTCACTCTGAGACACCAAAATTTCTTGACACCCGAA 1323
QY 841 GGGGAGTGGCTGTGACATTTGAGATTTGAGAGTGGAGAGAGAGAGAGAGAGAGAG 900
Db 1324 GGGGAGTGGCTGTGACATTTGAGATTTGAGAGTGGAGAGAGAGAGAGAGAGAGAG 1383
QY 901 TTCCAGAGCGTGGCTCTCTACTACCCCTGACAGAGAGAGAGAGAGAGAGAGAGATC 957
Db 1384 TTCCAGAGCGTGGCTCTCTACTACCCCTGACAGAGAGAGAGAGAGAGAGAGATC 1443
QY 958 TCTCTGACACCGGTGACACACAGATCCCTATGCTATGCTATGCTATGCTATGCTA 1017
Db 1444 TCTCTGACACCGGTGACACACAGATCCCTATGCTATGCTATGCTATGCTATGCTA 1503
QY 1018 GGGGCAAAAG 1077
Db 1504 GGGGCAAAAG 1563
QY 1078 CCGGCGACCTTCTCTACACACTGAA-----TGCCGATATAC 1116
Db 1564 CCGGCGACCTTCTCTACACACTGAAATGATGATGATGATGATGATGATGATGATGAT 1623
QY 1117 GAGTGTCTTACAG 1176
Db 1624 GAGTGTCTTACAG 1683
QY 1177 CATGAGGACCAACCATGCTGTGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1236
Db 1684 CATGAGGACCAACCATGCTGTGCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1743
QY 1237 GCCATCTGTGTATATTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1296
Db 1743 GCCATCTGTGTATATTTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1800

```

Db 1744 GCCATCCTGATATTCTGGAGGCACTTCCAGACACCATTAGTTGCTGGGCGG 1803
QY 1297 CCCATGCTTCCATGATGCTGACACTGCTGCTGGGATACATGATGATCCGCTGAC 1356
Db 1804 CCCATGCTTCCATGATGCTGACACTGCTGCTGGGATACATGATGATCCGCTGAC 1863
QY 1357 GTGGGGCCGCCAAGTCTCCACCTGCTGCTGCGCCAGGCGCTTCTTCCCTGCTTC 1416
Db 1864 GTGGGGCCGCCAAGTCTCCACCTGCTGCTGCGCCAGGCGCTTCTTCCCTGCTTC 1923
QY 1417 ACAATTTGATGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1476
Db 1924 ACAATTTGATGCTGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1983
QY 1477 GCCAGCGCTTCCACGCGCTACAGCTAGTGGGTCCGCTACACAGGGGCGCTTACGCTCT 1536
Db 1984 GCCAGCGCTTCCACGCGCTACAGCTAGTGGGTCCGCTACACAGGGGCGCTTACGCTCT 2043
QY 1537 ATGGCAATTTACAGGCTACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1594
Db 2044 ATGGCAATTTACAGGCTACAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2103
QY 1595 CTCAGTCCCAAC-CCGCTAGTGAACCCGATGACCCCAAGATGACATTTGCTGCTGAC 1653
Db 2104 CTCAGTCCCAAC-CCGCTAGTGAACCCGATGACCCCAAGATGACATTTGCTGCTGAC 2163
QY 1654 CCCAATACCGCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1713
Db 2164 CCCAATACCGCAACAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2223
QY 1714 GGTTCAGCTTCCGCTACATGAGGCAAGAGCTGCGCCACCACTACAAAGAGGCGCACTTC 1773
Db 2224 GGTTCAGCTTCCGCTACATGAGGCAAGAGCTGCGCCACCACTACAAAGAGGCGCACTTC 2283
QY 1774 ATCACTTCAGATGACCTTCTATTTACCTCATCGCTGCTGCTGCTGCTGCTGCTGCT 1833
Db 2284 ATCACTTCAGATGACCTTCTATTTACCTCATCGCTGCTGCTGCTGCTGCTGCTGCT 2343
QY 1834 GCCTACAGCGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1893
Db 2344 GCCTACAGCGGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2403
QY 1894 GCCATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1953
Db 2404 GCCATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2463
QY 1954 AACAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2520
Db 2464 AACAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-035-045-23
Query Match 96.0%; Score 1928.6; DB 9; Length 2520;
Best Local Similarity 98.2%; Pred. No. 0;
Matches 2001; Conservative 0; Mismatches 9; Indels 27; Gaps 4;

QY 1 ATCACTACAGCGCCATCAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
Db 484 ATCACTACAGCGCCATCAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 543
QY 61 CGTACCAACCCAGCGCCGACACACACACACACACACACACACACACACACACACACAC 120
Db 544 CGTACCAACCCAGCGCCGACACACACACACACACACACACACACACACACACACACAC 603
QY 121 CGCTGGAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 180
Db 604 CGCTGGAATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 663
QY 181 CTGCTTGGAGGCGCTGCGCCGCGGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 240
Db 664 CTGCTTGGAGGCGCTGCGCCGCGGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 723
QY 241 ACACTGAGCCCAACCAACATGATGCTGAGAGAGGCGCCAGCGCTGCTGCTGCTGCTG 300
Db 724 ACACTGAGCCCAACCAACATGATGCTGAGAGAGGCGCCAGCGCTGCTGCTGCTGCTG 783
QY 301 GACAAAGTGAAGAGACAGACAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 360
Db 784 GACAAAGTGAAGAGACAGACAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 843
QY 361 TACACTTCTTCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
Db 844 TACACTTCTTCAATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 903
QY 421 GAGTCTGGGCGATGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Db 904 GAGTCTGGGCGATGACCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 963
QY 481 TTCTGCGCATGACATGACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 540
Db 964 TTCTGCGCATGACATGACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1023
QY 541 GGCCACAGGCTGCGCCGCGCCACCCCTGAGAGACAGACAGAGCTTACCTGCAACAG 600
Db 1024 GGCCACAGGCTGCGCCGCGCCACCCCTGAGAGACAGACAGAGCTTACCTGCAACAG 1083
QY 601 GAGTGGCAACAATGCTGGAAGCGCACTTGTCTTCAACACATCTGACAGCTGCTG 660
Db 1084 GAGTGGCAACAATGCTGGAAGCGCACTTGTCTTCAACACATCTGACAGCTGCTG 1143
QY 661 GAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 720
Db 1144 GAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1203
QY 721 CTCTCGCTGTGCAAAACACCTGCAACAGAGGCTGCTTACCCCTGCAAGCTGCTT 780
Db 1204 CTCTCGCTGTGCAAAACACCTGCAACAGAGGCTGCTTACCCCTGCAAGCTGCTT 1263
QY 781 GAGGAGATCTGGAAGGCTGCAACTTCACTCTCTGAGACCAACAATCTTCTGACCGCA 840
Db 1264 GAGGAGATCTGGAAGGCTGCAACTTCACTCTCTGAGACCAACAATCTTCTGACCGCA 1323
QY 841 GGGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
Db 1324 GGGGAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1383
QY 901 TTCCAGAGCGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 957
Db 1384 TTCCAGAGCGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1443
QY 958 TCTCTGCAACCGCTCAACCAACAGATCCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1017

```

```

RESULT 5
US-10-035-045-23
; Sequence 23, Application us/10035045
; Publication No. US20030054448A1
GENERAL INFORMATION:
APPLICANT: ADLER, JON ELLIOT
APPLICANT: LI, XIAODONG
APPLICANT: STASZEWSKI, LENA
APPLICANT: O'CONNELL, SHAWN
APPLICANT: ZOZULYA, SERGEY
TITLE OF INVENTION: TIR TASTE RECEPTORS AND GENES ENCODING SAME
FILE REFERENCE: 078003-0280681
CURRENT FILING DATE: 2002-01-03
PRIOR FILING DATE: 2001-01-03
PRIOR FILING DATE: 2001-01-03
PRIOR FILING DATE: 2001-04-19
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 23
LENGTH: 2520

```


OY	598	CAGAGGAGCGACACACATGCGCCGAGACGGCACCTGTTGCTTCCATCAACCATCTCAGGCTGCT	657
Db	1105	CAGAGCTGTGACGCTCTGCTTGAACACCAACAGTCTCTTAACAAACATCTTATATTTTCG	1164
OY	658	GGGAGAGGCTGTCGTACTACAGCTGTACTCTGCGGCTATATGCTGTGGCCATGCGCCCTGCAC	717
Db	1165	GGGAGAGCGCGGAGGTCTACAGCTGTACTCGGAGGTTTACGGCGGTGGCCAGTCCCTCCAC	1224
OY	718	AGCCTCTCGGCTGTGACAAAACACCTGCACCCAGAGGGTGGTCTACCCCTGGCAGCTG	777
Db	1225	AGACTCTCGGCTGTAAACCGGGTCCGCTGCACCAACAAAAGGTCTACCCGTGGCAGCTA	1284
OY	778	CTTAGAGAGATCTGGAGAGTCACTTCACTCTCTGGACCAACAAATCTTCTTCGACCG	837
Db	1285	CTCAGGGAGATCTGGACAGCTCAACTTCACGCTCTGTGGTACACGGGCTCTTCTTACCAA	1344
OY	838	CAAGGGAGCTGGCTCTGCACCTTGGAGATTGTTCAGTGGCAATGGAGACCGAGCCAGAT	897
Db	1345	CAAGGGGACATGGCGAGTGTCTTGGACATCATCAGATGGCAGTGGGACTGTGAGCCAGAT	1404
OY	898	CCCTTCAGAGCGCGGCTGCTACTACCCCTGTGCAGCGCAGCTGATGAACAT---	954
Db	1405	CCCTTCAGAAAGCATGCGCTCTCTATTCTCCACAGCAGAGAGCTTAACCTTATACAT	1464
OY	955	ACATCTCTGCACACGCTCAACAACAGATCCCTATGTCTCATGTTCACAAGAGTGGCCAG	1014
Db	1465	GTGTCTCTGTATACACCCCAACACAGGTGCCGTGTCTCATGTGTCCAAAGCTGCGCAG	1524
OY	1015	TCAGAGCGAAAAGAAAGACCTGTGGGCTATCCAGCTGTGCTGTGCATGCATCGATGCG	1074
Db	1525	CCAGAGCGAAAATGAAGATGTGTGGGCGCTCCACCCCTTGTGGCTTGAGAGCTTGATGT	1584
OY	1075	CTTCCCGGCACTTCCCAACCACTGAA-----TGCCCAAT	1113
Db	1585	ATGCGAGGCACTTACCTCAACCCGCTCAGCAGATGATTAAGTGTCTGTCCGCGCCGGT	1644
OY	1114	AACGAGTGTGCTTACAGAGTGTGAGACTCCGTGTTCAACGCGCAGCTGGTGTCTCTGGA	1173
Db	1645	TGCATGTGTGCTTACAGAAACGACATACATGTGTTCCACAGGGCGGCTACATCTCTGAG	1704
OY	1174	TGCGATGAGSCACCCACCATGCTGTGTGGCCCTGCTGGCCGCTTGAGCTCTCAGCAC	1233
Db	1705	TGGCAGCAAGTGGCCACATCTGTGGTGGCCATCTGGTGGCCCTTGAGCTTCTTCAGTACA	1764
OY	1234	CTGGCAGATCGGGATATCTGTGGAGCAATTCACACACCATATGTTGCGCGGCTGAG	1293
Db	1765	CTGGCATCTTTTCATCTTCTGGAGCAATTTCCAAACACCCATGGTGGCTCGGCGCGT	1824
OY	1294	GGCCCCATGTGCTTCTGATGTGTGACACTGTGCTGTGGATGACATGGTGTCCCGGTG	1353
Db	1825	GGCCCCATGTGCTTCTGATGTGTGACACTGTGCTGTGGATGACATGGTGTCCCGGTG	1884
OY	1354	TACGTGGGGCGCCCAAGTGTCCACCTGCTGTGCGCCAGAGCCCTTGTCCCTGTGC	1413
Db	1885	TATGTGGGGCGCCCAAGTGTCTTCTGATGTGTGACACTGTGCTGTGGATGACATGGTGTCCCGGTG	1944
OY	1414	TTTCACATTTGCAATCTCTGTATTCGCGGCGGCTTTTCCAAATGTCTGCGGCTTTCAG	1473
Db	1945	TTTCTCATCTGCTATCTCATCTACCTGTGCGCTCTTCCAAATGTGTGTCTTTCAG	2004
OY	1474	ATGCGACAGCCGCTTCCACAGCGCTACAGCTTACGTGGGTCCGCTACAGGAGCGCTTACATC	1533
Db	2005	ATGCGACAGAGCGCTTCCCAAGTGTGCTTTCAGATGCTTTTGAATGCGTTACCAAGGCGCTTATGTC	2064
OY	1534	TCATATGCAATTTATCACGACTCAAAATGTCAATTTGTGTATATGGCACTGCTGGCAGG	1593
Db	2065	TTTCTGCGCTTCACTACAGGCACTCAAGTGTGGCCCTGTGTGTGGCAACATGCTGTGCAAC	2134
OY	1594	CCTCAATCCACCC---CGTACTGACCCCGATGAGACCCCAAGATCACAATGTCTCTCTGT	1650
Db	2135	ACCATCAACCCCATTTGGCGGAGCGACCGCGGATGAGCCCAACATCATGTATCTCTCTGTG	2184

QY	1651	AACCCCACTACCGCAACAGCCTGCTGTTTCAACACAGCAGCTGCTGCTCAGAG	1710
Db	2185	CACCCTAACACTCCGCAAGGGGCTACTGTTCAACACCAAGCATGGAATCTGCTGCTG	2244
QY	1711	GTGGGTTTCAGCTTGCCCTACATGAGGCAAAAGAGCTGCCACCACACTACACGAGGCCAAG	1770
Db	2245	CTGGGGTTTCAGCTTGCCCTACATGAGGCAAAAGAGCTGCCACCACACTACACGAGGCCAAG	2304
QY	1771	TTCATCACCCCTCAGATGATGACCTTTAATTTCACCTCATCGCTGCTGCTGACCTTCATG	1830
Db	2305	TTCATCACCCCTCAGATGATGACCTTTAATTTCACCTCATCGCTGCTGCTGACCTTCATG	2364
QY	1831	TCTGCCCTACAGCGGGGTGCTGTGTCACCATGCTGAGACCTCTTGGTCACTGTGCTCAACCTC	1890
Db	2365	TCTGTGCACAGAGCGGGGTGCTGTGTCACCATGCTGAGACCTCTTGGTCACTGTGCTCAACCTC	2424
QY	1891	CTGGCCATCACAGCCTGGGCTACTTTCGGCCCCAAGTCTCATATGATATGCTCTTCAACCGGAG	1950
Db	2425	CTGGCCATCACAGCCTGGGCTACTTTCGGCCCCAAGTCTCATATGATATGCTCTTCAACCGGAG	2484
QY	1951	CGCAACACAGCCCGCTACTTCAACAGCATGATCCAGGGCTACACCATGAGAGGAGCACTAG	2010
Db	2485	CGCAACACCTCAGCCTATTTCATATGACATGATCCAGGGCTACACCATGAGAGGAGCACTAG	2544
RESULT 7			
US-09-927-315-10			
; Sequence 10, Application US/09927315			
; Publication No. US20030040045A1			
GENERAL INFORMATION:			
APPLICANT: Zuker, Charles S.			
APPLICANT: Ryba, Nicholas J.P.			
APPLICANT: Nelson, Greg			
APPLICANT: Hoon, Mark A.			
APPLICANT: Chandrashekar, Jayaram			
APPLICANT: Zhang, Yifeng			
APPLICANT: The Regents of the University of California			
APPLICANT: The Government of the United States of America			
APPLICANT: as represented by the Secretary of the			
APPLICANT: Department of Health and Human Services			
TITLE OF INVENTION: Mammalian Sweet Taste Receptors			
FILE REFERENCE: 02307E-120110US			
CURRENT APPLICATION NUMBER: US/09/927,315			
CURRENT FILING DATE: 2001-08-10			
PRIOR APPLICATION NUMBER: US 60/302,898			
PRIOR FILING DATE: 2001-07-03			
NUMBER OF SEQ ID NOS: 25			
SOFTWARE: PatentIn Ver. 2.1			
SEQ ID NO 10			
LENGTH: 2993			
TYPE: DNA			
ORGANISM: Rattus sp.			
FEATURE:			
OTHER INFORMATION: rat T1R2 sweet taste receptor			
US-09-927-315-10			
Query Match	61.3%;	Score 1232.8;	DB 9; Length 2993;
Best Local Similarity	76.9%;	Pred. No. 0;	
Matches 1568; Conservative	0;	Mismatches 442;	Indels 30; Gaps 4.
QY	1	ATACCTTACAGGCGCATCAGCATGAGCTGCGAGACAAAGTGCTGCCGGCTTGGCTG	60
Db	505	ATTCATCATNAGGGCCATCTCCGACAAAGCTGGGGGACAAAGCGGCACACTTCCCTGACATGCTA	564
QY	61	CGTACACACACCCAGCGCGACCAACACAGTCGAGGCGCATGCTGACGTATGCTGCACCTTC	120
Db	565	CGCACAGTGTCCAGCGGCCACCCACACACATCGAGGCGCATGCTGACGTATGCTGACCTTC	624
QY	121	CGCTGGAATCGATCATGTGCTGTGGTGGAGCAGCAGACACTATGGGCGGAGCAATGGCCAG	180
Db	625	CAATGGAATCGATGTGTGCTGTGTGAGCAGCAGACGATTAAGCGCCGCGAGAACAGCCAC	684
QY	181	CTGCTTGGGAGCGGCTGGCC---CGGCGGACATCTGCAATGCGCTTCACGAGAGAGCTG	237

TYPE: DNA
ORGANISM: Rattus sp.
FEATURE:
OTHER INFORMATION: rat G-protein coupled receptor (GPCR) B4
OTHER INFORMATION: nucleotide sequence
US-10-124-598-3

Query Match 61.38; Score 1232.8; DB 12; Length 2993;
Best Local Similarity 76.98; Pred. No. 0;
Matches 1568; Conservative 0; Mismatches 442; Indels 30; Gaps 4;

QY 1 ATCACTACAGAGCCATCAGATGAGTGGAGACAAGTCCGCTTCCGCTTGGTG 60
DB 505 ATCACTACAGAGCCATCAGATGAGTGGAGACAAGTCCGCTTCCGCTTGGTG 564
QY 61 CGTACCAACAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 120
DB 565 CGCAGAGTGGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 624
QY 121 CGCTGGAAGTGGATCATTTGCTGTGTGAGCAGCAGCACTATGGCCGCGAGAGCCAG 180
DB 625 CAATGGAAGTGGATCATTTGCTGTGTGAGCAGCAGCACTATGGCCGCGAGAGAGCCAG 684
QY 181 CTGCTGGCGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 237
DB 685 CTGCTGGCGAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 744
QY 238 CCCACACTGACCCCAACAGATGACGTCAGAGGAGCCGAGCCGCTGCTGCTGCTGCTG 297
DB 745 CCCATACCTGAGTCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCCAGCC 804
QY 298 GTGACAGAGTGTGACAGAGCAGCAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 357
DB 805 CTGACAGAGTGTGACAGAGCAGCAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 864
QY 358 CTGACAGAGTGTGACAGAGCAGCAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 417
DB 865 CTGACAGAGTGTGACAGAGCAGCAGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 924
QY 418 TCCGAGTCTGGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 477
DB 925 TCTGAGTCTGGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 984
QY 478 ACCTTCTGGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 537
DB 985 ACCTTCTGGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1044
QY 538 TGGGAGCCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 597
DB 1045 CGCCGTACAGAGCCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1104
QY 598 CAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 657
DB 1105 CAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1164
QY 658 GGGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 717
DB 1165 GGGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1224
QY 718 AGCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 777
DB 1225 AGCTCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1284
QY 778 CTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 837
DB 1285 CTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1344
QY 838 CAGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 897
DB 1345 CAGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1404
QY 898 CCCTTCCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 954
DB 1405 CCCTTCCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 954

DB 1405 CCCTTCCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1464
QY 955 ACATCTGTGACAGCCAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTGACAGTGTG 1014
DB 1465 GTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1524
QY 1015 TCAGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1074
DB 1525 TCAGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1584
QY 1075 CTGCTGGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1113
DB 1585 CTGCTGGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1644
QY 1114 AACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1173
DB 1645 AACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1704
QY 1174 TGGCATGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1233
DB 1705 TGGCATGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1764
QY 1234 CTGCGCATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1293
DB 1765 CTGCGCATCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1824
QY 1294 GGGCCATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1353
DB 1825 GGGCCATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1884
QY 1354 TACGTGGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1413
DB 1885 TACGTGGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1944
QY 1414 TTGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1473
DB 1945 TTGCAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2004
QY 1474 ATGAGCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1533
DB 2005 ATGAGCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2064
QY 1534 TCTATGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1593
DB 2065 TCTATGATTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2124
QY 1594 CCGTACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1650
DB 2125 CCGTACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2184
QY 1651 AACCCAACTACGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1710
DB 2185 AACCCAACTACGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2244
QY 1711 GTGGGTTTACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1770
DB 2245 GTGGGTTTACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2304
QY 1771 TTGATCACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1830
DB 2305 TTGATCACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2364
QY 1831 TCTGCTACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 1890
DB 2365 TCTGCTACAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2424
QY 1891 CTGGCCATCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1950
DB 2425 CTGGCCATCAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2484
QY 1951 CGCAACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2010
DB 2485 CGCAACAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2544

RESULT 9
US-10-096-144-4

; Sequence 4, Application US/10096144

Publication No. US200300222288A1

GENERAL INFORMATION:

APPLICANT: Adler, Tony Ellis

APPLICANT: Indemeter, Inc.

APPLICANT: The Regents of the

1 TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein

; TITLE OF INVENTION: Involved in Sensory Transduction

FILE REFERENCE: 02307E-088720US

CURRENT APPLICATION NUMBER: 05/10/096,144

; CURRENT FILING DATE: 2002-03-08
 ; PRIOR APPLICATION NUMBER: 00/261 631

PRIOR FILING DATE: 1999-07-27
 PRIOR APPLICATION NUMBER: 09/301,031

PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER:

PRIOR FILING DATE: 1998-07-28

; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER

PRIOR FILING DATE: 1998-12-17

NUMBER OF SEQ ID NOS: 10

```
;; SOFTWARE: PATENTLIN Ver. 2.0
;; SEQ ID NO 4
```

LENGTH: 2532

TYPE: DNA

ORGANISM: Mus sp.

FEATURE:

OTHER INFORMATION:

OTHER INFORMATION: nucleotide sequence
MS-10-095-144-4

44-38861-107-50

Query Match	60.3%	Score 1212.6	DB 9	Length 2532
Best Local Similarity	76.3%	Pred. No. 0		
Matches 1555; Conservative	0	Mismatches 454	Indels 30	Gaps 4

OY	2	TCACCTACAGGCGCATATGAGTATGAGCTGGAGACAAAGGTGGGGCTTCCGGCTTTGGCG	61
Db	494	TCACATATAGCGGCATATACCGACAAAGCTGCMAACAGCGGGCTTCCCTGCACTATCG	553
OY	62	GTACCAACCCAGGCGCGGACCAACACGTGAGGCGCATGGTGTGACGTATGTGCACCTTC	121
Db	554	GCACTGTGCCAGGCGCCACCACACATATGAGGCGCATGGTGTGCAACTGATGTTCACCTTC	613
OY	122	GCTGGAAGTGGATATATGTGCTGTGTGAGAGGACACACATATGGCGCGACATATGGCCAGC	181
Db	614	AGTGAAGCTGGATCTGTGTCTGTGTGAGCATATGACATTTATGGCCGAGGAACACCCACC	673
OY	182	TGCTTGGCGAGCGGCTGGC---CCGGCGGCACATCTGCATCGCTCCCTTCAGAGACGTGC	238
Db	674	TGCTGAGCCAGCGCTGATCCACACATGCGGCACATCTGCATTTGCTTCAGAGAGTTCTGC	733
OY	239	CCACATCGAGCCCAACCAACATATACGTGCAGAGAGAGGCCACGCCCTGGTGACATTTG	298
Db	734	CCGTACCAAGAACCCAAACACAGCTGTGTAGGCGCTTAGGAGAGGAGACCAACTGTGCACCAATCC	793
OY	299	TGGAACAAGCTGCACAGACAGCACAGCGCGCTCTGTGCTCTGTCTTCGCCCGACTGCACC	358
Db	794	TGGAACAAGCTGCGGGCGGACCTTGGGCGCGTGTGTGTGATTTCTTGCCGGAGATCGAGCC	853
OY	359	TGTACCACTTCTTAATGAGGTGCTGCGCCAAACTTCACGGGCGCGCTGTGTGATCGCT	418
Db	854	TGCACAACCTTCTCCGAGAGTCTGCGGTGGAACCTTCACGGGCTTTGTGTGAATTGCT	913
OY	419	CCGAGTCTCTGGGCATGAGCCCGGCTCCCTGCACAACCTCACAGGATGGGCGCACTTGGGA	478
Db	914	CTGAGTCTCTGGGCATGAGCCCTGTCTTCTAACACAACCTCACAGAGCTGCCACACAGGGA	973
OY	479	CTTCTCTGGGCATACCATCCAGAGGCTGCCATCCCGGGCTTCACTGATGTTCGGGAGT	538
Db	974	CTTCTCTGGGCTTACCATCCAGAGGGTGTCTCATCTCTGTGCTTCAACCAAGTTCAGAGTC	1033

QY	539	GGGGCCCAAGCGTGGGGCCGCCACCCCTAGCAGAGACCAAGCCCAAGCTATACCTGCACAC	598
Db	1034	GCATGACAAAGCCAGGGGTATCGATGCTCTACAGACAGCGCTGGAGCTACTCTGTAAAC	1093
QY	599	AGGATGTCGACAACTGGCTCTGAAGCCACCGTCTGTCCTTCAACACATTTCTACAGGCTCTTG	658
Db	1094	AGGACTCGACGCGCTGCATATACATACATGAGTCCCTTCAACAACGTTCTCACTGCTTTCGG	1153
QY	659	GGGAGCGTGTGCTTACAGCGGTGTACTGTGGCTGTATGCTGTGGCCATVGGCCCTGCACA	718
Db	1154	GGGAGCGTGTGCTTACAGCGGTGTACTGTGGCCCTGTACAGCGGTGGCCCAACACCTTCACA	1213
QY	719	GCCTCTCGGCTGTGACAAAAGCACTGCACCAAGAGGGTGGTCTTACCCCTGGCAGCTGC	778
Db	1214	GACTCTCTCACTGTGAATTCAGGTCCGGTGCACCAAGCAAAATCTGTATCATATGGCAGCTAC	1273
QY	779	TTGAGGAGATCTGCAAGTCACTCACTCACTCCCTGGACCAACAAATCTTCTTCGACCGC	838
Db	1274	TCAGGAGATCTGCAAGTCACTCACTCACTCCCTGGGCAACCAAGCTCTTCTTCGACGAC	1333
QY	839	AAGGGAGCTGGCTCTGCACTTGGAGATTGTCCAGTGGCAATGGGACGGAGCCGAAATC	898
Db	1334	AAGGGAGATCTGCACTCTCTCTGGACATCATTCAGTGGCAGTGGGAGCTGCAGCCAGAAC	1393
QY	899	CCTTCACAGAGGTGGCTCTCTACTACCCCTGCAGGACGACGCTGAAGAAGAT--CAAGA	955
Db	1394	CCTTCACAAAGATAGCTCTCTACTCCCCACCGACGAGGGGTGACCTACATTAGCAATG	1453
QY	956	CATCTCGACACCGCTCAACACACGATCCCTCTATGCATGTGTTCACAAGAGTCCAGT	1015
Db	1454	TGTCTGTGTACACCCCAACACACGCTCCCATATCATGTGTCTTAAGAGTTGCCAGC	1513
QY	1016	CAGGCAAAAGAAAGAACCTGTGGGATCCACGCTGCTGCTTCGAGTCACTGCATGCACC	1075
Db	1514	CTGGCAAAATGAAAAAACCAATFAGGCGCTCCACCACTGCTGCTCGAGTGTGGAGCTGTC	1573
QY	1076	TTCCCGGACCTTCTCTAACCAACATGAA-----TGCCCGAATA	1114
Db	1574	CGCCGGACACTTACTCAACCGATACAGTAAATGATTAACTGTGTCTGCTCCCGGGATT	1633
QY	1115	ACGAGTGTCTCTACAGAGTGAAGACGTCCTGCTTCAAGGCGGACGTCGTCCTCGAAT	1174
Db	1634	CCATGTGTTCTTACAAABAACAACATGCTTGGCTTCAAGGCGGCGCTGCTCTTCTTGAGT	1693
QY	1175	GGCATGAGGACCCACCACTGCTGTGGCCCTGCTGGCCGCCCTGGGGCTTCTCAAGACCC	1234
Db	1694	GGCAAGAGTCCCACTATGCTGGTGAACATCTCTGGCCGCCCTGGGGCTTCACTCAATAGC	1753
QY	1235	TGGCCATCTGTGTATATTCTGAGAGCACTTCAGACACCCATAGTGGCTGGCTGGGG	1294
Db	1754	TGGCCATCTGCTCATCTTCTTGGAGCAATTTCCAGAGGCCCAATGGTGCCTGTGGGCGAG	1813
QY	1295	GCCCATGTGCTTCTCATGTGTGACACTGTGTGTGGCATACATAGGAGGCGCCGGGT	1354
Db	1814	GCCCATGTGCTTCTCATGTGTGTGGCTCTGCTGTGGCTGTGGAGTGTGCCCGGT	1873
QY	1355	ACGTGGGGCCGCCAAGGTCTCCACCTGCTCTGCCCGCAGGCGCTCTTCCCTCTGCT	1414
Db	1874	ATGTGGGCCCCCGCAGGCTCTTCCCTGTTCTTGGCCGAGGCTTCTTCACAGTTTGGT	1933
QY	1415	TCACAAATTGCATCTCTGTATGCGCGGTGCTTCTTCCAGATGCTGTGCGCTTCAAAGA	1474
Db	1934	TCCTCGTCTGCTCTCTCTCGATATACAGGTGGCGCTCTTCCAGATTGTGTGCTGCTTCAAGA	1993
QY	1475	TGGCCAGCGGCTTCCCAAGGCGCTACAGCTACAGTGGGTTCGCGTACACAGGGGCGCTACGCT	1534
Db	1994	TGGCCAGCGGCTTCCCAAGGCGCTACAGGTTTCTGAGTGTGTTACACAGGGGCGCTACGCT	2053
QY	1535	CTATGCAATTTATACAGTACTCAAAATGCTCAATTTGAGTAATTTGCACTGTGCAGCGC	1594
Db	2054	TCGTGGCTTATATACGGGCGCTCAAGTGGCCCTGTGGTGGGGAACAATGCTGGCGCACCA	2113
QY	1595	CTCAATCCACACC--CCGATACGACCCCGATGACCCCAAGATCAACAATGTGTCTCTGTA	1651

QY	1235	TGGGCATCTCGGAGATATTCCTGGAGGCACTCCAGACACCCATATTCGCTCGCTGGG	1236	
Db	1754	TGGGCATCTCTGCTCATCTCTCTGGAGCATTTTCCAGACGCCATAGTGTGGCTGGGGGCG	1813	
QY	1295	GGCCCATGTCTCTTCGATGCTGATGCTGATGCTGCTGCTGGTGGCATACATAGGAGTCCGGTGT	1354	
Db	1814	GGCCCATGTGCTTCTGTGATGCTGTGTCGCCCTGTGCTGTGGCGATTGGATGCCCGGT	1873	
QY	1355	ACGTGGGGCGCCCAAGGTCTCCACTGCTTGGCCGAGGCCCTCTTTCCTCTGCT	1414	
Db	1874	ATGTGGGGCCCCCAGGCTTCTCTCTGTTTGTGGCCGAGGCTTCTTTCACCGTTTGTCT	1933	
QY	1415	TCACAATTTGATATCTCTCTGTATGCGCGGTGTTCTTCCAGATGCTGTGCGCTTCAGA	1474	
Db	1934	TCTCTGCTGTGCTCTCTCTGCTGATACGGTGGCGCTCTTCCAGATGTTGTGCGCTTCAGA	1993	
QY	1475	TGGCCAGCGGCTTCCAGCGGCTTACAGCTACTGAGTCCGCTGCTACAGAGGGGCTCTAGCT	1534	
Db	1994	TGGCCAGAGCGCTGCGCAAGGCGCTTACGGTGTTCGTATGGGTATACACGGGGCTCTAGCT	2055	
QY	1535	CTATGGCATTTATCAGGGTACTTAAATAGTCTATTGTGTAAATTGGCATGCTGGCAGCGC	1594	
Db	2054	TCTGTGGCTTATCAGCGGCGGCTGTAAGGTGGCCCTGTGTGGGGAACAATGCTGGGCAACA	2113	
QY	1595	CTCAGTCCCAACC---CCGATATGACCCCGATGACCCCAAGATTCAAATGTCTCTGTA	1651	
Db	2114	CCATCAACCCCATATGGCGCGGACCGACCCGAGTACCCCATATTCATTAATCCTCTCTGCC	2173	
QY	1652	ACCCCAACTACCGCAACAGGCTGCTGTTCAACACAGCCGTGAGACCTGCTGCTCTAGTGG	1711	
Db	2174	ACCTCAACTACCGCAACGGGCTACTCTTCCACACAGCATGAGATCTGTGCTGTCCGTCG	2233	
QY	1712	TGGGTTTACACTTCGCTTACATGGGCAAGAGCTGCCCCACCACTACATCAACGAGGCCAAGT	1771	
Db	2234	TGGGTTTACACTTCGCTTACATGGGCAAGAGCTGCCCCACCACTACATCAACGAGGCCAAGT	2293	
QY	1772	TCACTACCCCTCAGATGAGACCTTGTATTTTACGCATGCGTCGTCGCGACCTTCAAGT	1831	
Db	2294	TCACTACCCCTCAGATGAGACCTTGTATTTTACGCATGCGTCGTCGCGACCTTCAAGT	2355	
QY	1832	CTGCGTACAGCGGGGCTGCTGATCACCATGCTGAGACCTTGTGCTCACTGTGTCAACCTTCC	1891	
Db	2354	CTGCGTACAGCGGGGCTGCTGATCACCATGCTGAGATCTCTGTGCTCACTGTGTCAACCTTCC	2413	
QY	1892	TGGGCATCTACGCTTGGGCTACTTGGGCCCCAAGTGTCTATGATCTCTTCTTACCCGGAGC	1951	
Db	2414	TGGGCATCTACGCTTGGGCTACTTGGGCCCCAAGTGTATGATCTCTTCTTACCCGGAGC	2473	
QY	1952	GCAACAGCGCCGCTTACTTAAACAGATGATCCAGGGCTACACGATGAGGAGGAGCTAG	2010	
Db	2474	GCAACAGCTTACGCTTATTCTTAATAGATGATTCAGGGCTACACGATGAGGAGGAGCTAG	2532	
RESULT 11				
US-10-124-598-4				
Sequence 4, Application US/10124598				
Patent No. US20020119526A1				
GENERAL INFORMATION:				
APPLICANT: Zuker, Charles S.				
APPLICANT: Adler, Jon Elliot				
APPLICANT: Lindemeier, Juergen				
APPLICANT: The Regents of the University of California				
TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor				
FILE REFERENCE: 02307E-08872005				
CURRENT APPLICATION NUMBER: US/10124-598				
PRIORITY FILING DATE: 2002-04-16				
PRIORITY APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US/09/361,631				
PRIORITY FILING DATE: EARLIER FILING DATE: 1999-07-27				
PRIORITY APPLICATION NUMBER: EARLIER APPLICATION NUMBER: US 60/112,747				
PRIORITY FILING DATE: EARLIER FILING DATE: 1998-12-17				
NUMBER OF SEQ ID NOS: 10				

QY 899 CCTTCAGAGGCTGCTCTACTACCCCTGACGACAGCTGAAGACAT---CAAGA 955
 DB 1394 CCTTCAGAGGCTGCTCTACTACCCCTGACGACAGGCTGACCTGACCTTACGATG 1453
 QY 956 CATCTGACACAGCGTCAACACAGATCCCTATGCTGATGTTCAAGAGTCCAGT 1015
 DB 1454 TGCTCTGATACCCCAACACAGGCTCCCATATCCATGTTGTTAAGAGTTGCCAGC 1513
 QY 1016 CAGGGCAAGAAAGAGAGCTGTTGGGATCCAGCTGCTGTTGTTGATGATGATGCC 1075
 DB 1514 CTGGGCAAGTGAAGAAAGAGAGCTGATAGGCTCCAGCCATGCTGATGATGATGATG 1573
 QY 1076 TTCCGGACCTTCTCAACCCACATGTA-----TGCCGGAATA 1114
 DB 1574 CGCGGACACTACTCTCAACCCGATGATGATGATGATGATGATGATGATGATGATG 1633
 QY 1115 ACGAGTGTCTTACAGAGATGAGCTGCTGCTTCAAGCGGACAGCTGCTTCTGAT 1174
 DB 1634 CCATGTGTCTTACAGAGATGAGCTGCTGCTTCAAGCGGACAGCTGCTGCTGAT 1693
 QY 1175 GGCATGAGAGCACCACATGCTGTTGGGCTGTTGGGCTGTTGGGCTGTTGGGCTG 1234
 DB 1694 GGCATGAGAGCACCACATGCTGTTGGGCTGTTGGGCTGTTGGGCTGTTGGGCTG 1733
 QY 1235 TGCCATCTGCTGATATTTGAGAGCATTTCAGACACCCATGATGCTGCTGCTG 1294
 DB 1754 TGCCATCTGCTGATATTTGAGAGCATTTCAGACACCCATGATGCTGCTGCTG 1813
 QY 1295 GCGCATGCTGCTGATATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1354
 DB 1814 GCGCATGCTGCTGATATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1873
 QY 1355 ACGTGGGCGCCCAAGAGTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1414
 DB 1874 ATGTGGGCGCCCAAGAGTCTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1933
 QY 1415 TCACAAATTTGATCTCTGATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1474
 DB 1934 TCCTCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1993
 QY 1475 TGCGAGCGGCTTCCAGCGGCTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1534
 DB 1994 TGCGAGCGGCTTCCAGCGGCTTCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2053
 QY 1535 CTATGCAATTTATCAGGCTTCAAAATGCTGATGCTGATGCTGATGCTGCTGCTG 1594
 DB 2054 TCGTGGGCTTATCAGGCTTCAAAATGCTGATGCTGATGCTGATGCTGCTGCTG 2113
 QY 1595 CTCAGTCCACCC---CCGTAAGTACCCCGATGACCCCAAGATCAAAATTTGCTGCTG 1651
 DB 2114 CCATCAACCCCATTTGGCGGAGCCGACCCCGATGACCCCAATATATATATCTCTGCTG 2173
 QY 1652 ACCCACTACCCGACAGAGCTGCTGTTCAACACAGGCTGCTGCTGCTGCTGCTGCTG 1711
 DB 2174 ACCCACTACCCGACAGAGCTGCTGTTCAACACAGGCTGCTGCTGCTGCTGCTG 2233
 QY 1712 TGCGTTCAGCTTCCATACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1771
 DB 2234 TGCGTTCAGCTTCCATACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2293
 QY 1772 TCATCACCTTCAGATGATGCTTCTATTTACCTGCTGCTGCTGCTGCTGCTGCTG 1831
 DB 2294 TCATCACCTTCAGATGATGCTTCTATTTACCTGCTGCTGCTGCTGCTGCTGCTG 2353
 QY 1832 CTGCTTCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1891
 DB 2354 CTGCTTCAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2413
 QY 1892 TGCGCATCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1951
 DB 2414 TGCGCATCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2473

QY 1952 GCAACAGCCCGCTACTCTCAACAGCATGATCCAGGCTACACCATGAGGAGGACTAG 2010
 DB 2474 GCAACACTTACGCTTATTCTCAATGATGATGATGATGATGATGATGATGATGATG 2532
 RESULT 12
 US-09-361-652-4
 ; Sequence 4, Application US/09361652
 ; Publication No. US20030036630A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Zuker, Charles S.
 ; APPLICANT: Adler, Jon Elliot
 ; APPLICANT: Lindemeyer, Juergen
 ; APPLICANT: Ryba, Nick
 ; APPLICANT: Hoon, Mark
 ; APPLICANT: The Regents of the University of California
 ; TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
 ; FILE REFERENCE: 02307E-08861005
 ; CURRENT APPLICATION NUMBER: US/09/361,652
 ; EARLIER FILING DATE: 1999-07-27
 ; EARLIER APPLICATION NUMBER: US 60/094,465
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: Patentln Ver. 2.1
 ; SEQ ID NO 4
 ; LENGTH: 2771
 ; TYPE: DNA
 ; ORGANISM: Rattus sp.
 ; FEATURE:
 ; OTHER INFORMATION: rat G-protein coupled receptor B3 (GPCR-B3)
 ; US-09-361-652-4
 Query Match 19.5%; Score 392.8; DB 9; Length 2771;
 Best local Similarity 51.4%; Pred. No. 4,5e-101;
 Matches 1046; Conservative 0; Mismatches 942; Indels 48; Gaps 4;
 QY 2 TCACCTACAGCGCCATCAGAGGATGCTGCGAGCAAGTGTGCTTCCGCGTTTGTGCTG 61
 DB 553 TCACCTATGAGGCAAGAGCGGTGTACTAGTGCACAGCCAAAGTTCCTCTTCTTCTG 612
 QY 62 GTACACACAGCGCGGACACAGCTGAGAGCTGAGAGCTGAGAGCTGAGAGCTGAG 121
 DB 613 GTACCGTCCCGAGAGCGGACAGAGTGTGAGAGCTGAGAGCTGAGAGCTGAGAG 672
 QY 122 GCTGGAATGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 181
 DB 673 GGTGGGTGTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 732
 QY 182 TGTGTGCGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 241
 DB 733 CGTGGAGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 790
 QY 242 CACTGCAAGCCCAACCAAGATGACGTGAGAGAGCGGAGCGGCTGAGCAATGTGG 301
 DB 791 -----TTTCTCTGCGCGGTGGGTGATCCCGAGATGCAAGATGATGCTG 834
 QY 302 ACAAGCTGAGAGAGAGAGAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361
 DB 835 AGCATGTGGCTGAGAGAGAGAGAGCGGTGTGTGTGTGTGTGTGTGTGTGTGTGT 894
 QY 362 ACCACTTCTTCAATGAGTGTGCTGCGGCAAGCTTACAGGCGCGGCTGCTGCTGCTG 421
 DB 895 GAGGTCTCTCAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 954
 QY 422 AGTCTGAGGCTGAG 481
 DB 955 AAGATGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1014
 QY 482 TCCGAGGATCAGATCAG 541
 DB 1015 TGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1074

Oy 542 GCCCAGAGCTGGCCGCCACCCCTCAGACAGACCAGCCAGACTATACCTGC-----A 595
 Db 1075 ATGTACAGGGCGGTAAACAGCTGTGTCAGCGCTTGGCCGAGGGGCTCCGTGTACAGACTA 1134
 Oy 586 ACCAGAGCTGGACACATGCTCTGAAGCCACCTTGTCTTCAACACCATCTCAGGCTCT 655
 Db 1135 ACCAGCTGTGGCGGGAGTGGCCACACGTTACAGACTCTGTACATGCTCCACGCTTGAAGCT 1194
 Oy 656 CTGGGAGAGCTGTGCTTACAGAGCTGTACTGCGGCTGTATGCTGTGGCCATGCTCCGTC 715
 Db 1195 TCTCCATAGTGTGCGCTACAGAGTGTATGAGGCTGTGTAGCGCTGTGGCCACGGCTCC 1254
 Oy 716 ACAGCTCTCTGGCTGTGTACAAAGACACCTGCACCAAGAGGCTGTACCCCTGGCAGC 775
 Db 1255 ACCAGCTCTGGAGTGTACTTCTGTAGATCTGTGTCCAGAGCCCACTACCTCTGGCAGC 1314
 Oy 776 TGTGTAGAGATCTGGAAGGTCAACTTCTCTCTGAGCCACCAAACTCTTTCGACC 835
 Db 1315 TTTCTCAGCAGATCTACAAAGGTGATTTTCTTCTACATGAGAACTACTGTGGCATTTGATG 1374
 Oy 836 CGAAGGGAGCTGTGCTGTGACCTGTGAGATTTGTCCACTGTGCAATGGGACGGAGCCAGA 895
 Db 1375 ACAACGGGAGACTCTAGTGTACTACAGATCTATGCTGCTGGAGCTGGAATGGACTGTAT 1434
 Oy 896 ATCCCTTCAGAGCTGTGCT 955
 Db 1435 GGACCTTTGAGATCTATCTGTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1494
 Oy 956 CATCTCT-----GCACACCGTCAACAACAGATCTCTATGTCTCATGTGTTCAGAGGT 1009
 Db 1495 CAAAATCTCAGTGTGCGACGGGAAAGAACTACAGTGTCTGTGTCTGTGTGTGTGTGTGTGT 1554
 Oy 1010 GCCAGTCAGGCAAAAGAAAGAGCTGTGGGCAATCCAGCTGTGCTGTGCTGTGCTGTGCTGTG 1069
 Db 1555 GTCTGTGCGAGGACACAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1614
 Oy 1070 ACTGCTTCCCGGACCT 1111
 Db 1615 CTTGGAAAGCTGTGGACCTTCT 1674
 Oy 1112 ATTAAGATGTGTCTTACAGAGTGTGAGCTCTGTCTTCAAGGCGAGCTGTCTCTCTCTCT 1171
 Db 1675 CAGAAGATGTGGGACCCCAAGAGAGAGACTGTGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1734
 Oy 1172 AATGAGATGAGGACCCACCATCTGTGTGGCCCTGTGGCCCTGTGGCCCTGTCTCTCTCT 1231
 Db 1735 CTGTGCAATGAAACCATCTCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1794
 Oy 1232 CCTGTGCAATCTGTGTATTTCTGTGAGGACACTTCCAGACCCATAGTGTCTGTGGCTG 1291
 Db 1795 TGTGTGGAGCTGTGGCTGT 1854
 Oy 1292 GGGGCGCATGTGTCTGT 1351
 Db 1855 GGGGAGGCTGTGTCTGT 1914
 Oy 1352 TGTAGCTGGGCGCCCAAGGTCTCTACCTGTCTGCGGACGAGGCGCTTCTCTCTCTCT 1411
 Db 1915 GCTTCTTGGGGAGCCCAAGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1974
 Oy 1412 GCTTCAATTTGATCTCTCTGTATGCGCGTGTGTCTTCTCTCTCTCTCTCTCTCTCTCT 1471
 Db 1975 GGTGTGCAATCTTCT 2034
 Oy 1472 AGATGGCCAGCGCTTCT 1531
 Db 2035 AATTTTCTACCAAGGTGTCCCAATCTCTACCTGTGGGCCCAAAACCAATGTGTGAGGTCT 2094
 Oy 1532 TCTCTATGGCATTTTACAGAGTACTCAAAATGTGTATTTGTGTGTGTGTGTGTGTGTGT 1591
 Db 2095 TATTTGTCTATTTGTCAAGTCTCAAGGTCTCAATTTGTCTCTCTCTCTCTCTCTCTCT 2154
 Oy 1592 GGCCTCAGTCCACCCCGTACTGACCCCGATGACCCCAAGATCACAATTTGTCTCTCTGTA 1651

Db 2155 GGACCCCAAGACCCCAAGGGAATATACAGCGCTTCCCATCTGTGTGTGTGTGTGTGTGTGA 2214
 Oy 1652 ACCCACTACCCGCAACAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1711
 Db 2215 CAGAGGTCAACTCTGTAGGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2274
 Oy 1712 TGGTTTACGCTTGTGCTTACATGAGGCAAGAGCTGTGCGCCCAACCAATCAAGCGGCAAGT 1771
 Db 2275 GTACTGT 2334
 Oy 1772 TCAATACCCCTGACATGTGATCTTGTATTTTCACTTCACTTGTCTCTCTCTCTCTCTCTCT 1831
 Db 2335 GTGTACCTGTAGCGCTGT 2394
 Oy 1832 CTGCTTACAGCGGGGT 1891
 Db 2395 GCATTTACAGGAGGCTACCTGTGCGGTCAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2454
 Oy 1892 TGGCATCAGCGCTGT 1951
 Db 2455 GCGGCGCTTACAGCGGT 2514
 Oy 1952 GCACACGCGCCGCTACTTCAACAGCATGATTCAGAGGCTTACACATGAGAGGGGAC 2007
 Db 2515 TCAACATATACACACACTTTCAGGCGCTCATGTCAAGAGACTACAGAGGCGCTGCGGG 2570

RESULT 13

US-09-927-315-4
 ; Sequence 4, Application US/09927315
 ; Publication No. US20030040045A1
 GENERAL INFORMATION:
 ; APPLICANT: Zuker, Charles S.
 ; APPLICANT: Ryba, Nicholas J.P.
 ; APPLICANT: Nelson, Greg
 ; APPLICANT: Hoon, Mark A.
 ; APPLICANT: Chandrasekar, Jayaram
 ; APPLICANT: Zhang, Yifeng
 ; APPLICANT: The Regents of the University of California
 ; APPLICANT: The Government of the United States of America
 ; APPLICANT: as represented by the Secretary of the
 ; TITLE OF INVENTION: Mammalian Sweet Taste Receptors
 ; FILE REFERENCE: 02307E-120110US
 ; CURRENT APPLICATION NUMBER: US/09/927,315
 ; CURRENT FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: US 60/302,898
 ; PRIOR FILING DATE: 2001-07-03
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO: 4
 ; LENGTH: 2771
 ; TYPE: DNA
 ; ORGANISM: Rattus sp.
 ; FEATURE:
 ; OTHER INFORMATION: rat T1R1 sweet taste receptor
 US-09-927-315-4

Query Match 19.5%; Score 392.8; DB 9; Length 2771;
 Best Local Similarity 51.4%; Pred. No. 4.5e-101;
 Matches 1046; Conservative 0; Mismatches 942; Indels 48; Gaps 4;

Oy 2 TCACCTACAGCGCCATACAGATGAGCTGTGAGCAAGGTGGCTGCGGCTTGGCTCTGC 61
 Db 553 TCAGCTATGAGGCAACAGCGGTGTACTATGTGCCAAGGCAAGTCCCGCTTCTCTTC 612
 Oy 62 GTACACACCCAGCGCGGCAACAGCTGTGAGGCGCATGTGTGAGCTGTATGTCTCACTTCC 121
 Db 613 GTACCGTCCAGGTAGCGGCAACAGGTGTGATGTGTGAGCTGTGTGAGAGATTTTG 672
 Oy 122 GCTGGAATGTGATCTTGT 181

Db 673 GGTGGGTGTGATCTGCTCTATTGGCAGCTAAGGTATTTACGGGCGAGCTGGGTGTGACAG 732
 QY 182 TGTGTGGAGAGCGGTGGCCCGGCGACATCTGCATCGCTCTCCAGAGAGCGTGGCCA 241
 Db 733 CGCTGGAGAGACTGGCGGTGGCCCGGCGACATCTGCATCGCTCTCCAGAGAGACTGGCC-- 790
 QY 242 CACTGCAAGCCCAACAGAACTGACGTACAGAGAGCGGCGAGCGCTGGTGCATTTGTGG 301
 Db 791 -----TTTCTGCTCCGCGGTGGGTGACCCGAGAGATGACAGCATGATGATGC 834
 QY 302 ACAAGCTGCAGCAGAGAGAGCGGCGGTGGGTGGGTGGTGTGTGCTGCGCCGACCTGACCTGT 361
 Db 835 AGCATCTGGCTCAGGCGCAGAGACACCGGTGGTGGTGTGTGTCTTACCGGACCTGGCTGA 894
 QY 362 ACCACTTCTTCAATGAGTGTGCTGCGCAGAACTTACAGGGGCGCGGTGGTGTGATCGCTCG 421
 Db 895 GAGTGTCTTCTAGGTGCGTGGTGTGCTGGGCGAACCTGACCTGGGAAAGTGGGGTCCCTCAG 954
 QY 422 AGTCTGGGCGCATGACCCCGGTCTGTCACAACTCAGCGAGCTGGGCGACTTGGGCGACT 481
 Db 955 AAGACTGGGCGCATCTCCACAGTACATCACAGCGTGAAGTGGGATCCAAAGGCGATTGGGACGG 1014
 QY 482 TCTGGGCGATCCATCCAGAGCGGTGGCGCATCCGGGCGTTCAGTGGTTCGGCGAGTGGG 541
 Db 1015 TGTCTGGGTGGCGGTCTCCAGAGAGAGAGATCTCTGGCTGAGAGAGATTGAGAGTCTTT 1074
 QY 542 GCCCAGAGCGTGGGCGCGCACCCTCAGCAGAGAGAGCGCAGAGCTTACTGCTG-----A 595
 Db 1075 ATGTACAGGGGTGTAAAGCTGCTCCAGCGCTTCCCGGAGGGGTCCGTTGGTGGAGCACTA 1134
 QY 596 ACCAGAGGTGGAGACACTGCTGTGAAGCGCACCTTGTCTTACACCATTTCTCAGGCTCT 655
 Db 1135 ACCAGAGGTGGAGCT 1194
 QY 656 CTGGGAGCGGTGTGCTGACAGCTGATCTGCGGTCTATGCTGGGCGCGAGTGGCGCTG 715
 Db 1195 TCTCCATGAGAGCGCGCTACAGAGTGTATGAGGCTGTGTACGCTGTGGCCACGCGCTCC 1234
 QY 716 ACAGCGCTCTGCGGTGTGACAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 775
 Db 1235 ACCAGCTCTGCGGTGTGATCTGAGATCTGTGTCAGAGAGAGAGAGAGAGAGAGAGAGAGCT 1314
 QY 776 TGTCTGAGAGATGTGAGAGTCACTTCACTCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGCT 835
 Db 1315 TTTCTGACAGATCTAAAGTGAATTTCTTCTACATGAGAAATCTGTGGCATTTGTG 1374
 QY 836 CGAAGGAGAGAGTGTGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 895
 Db 1375 ACAAGGAG 1434
 QY 896 ATCCCTTCAG 955
 Db 1435 GAGACCTTTGAGATCATGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1494
 QY 956 CATCTCT-----GCACACGCTCAACAACAGAGATCCATATGCTCATGTGTCCAGAGAGT 1009
 Db 1495 CAAAG 1554
 QY 1010 GCCAGTGAAGGAG 1069
 Db 1555 GTCTGGAG 1614
 QY 1070 ACTGCTCTCCGCGACCTCTCTCAACCACTGAA-----TGCCCGA 1111
 Db 1615 CCTGCGAAGCTGGGAGCTTTCTCTCAACATGATGATGATGATGATGATGATGATGATGATGAT 1674
 QY 1112 ATATAGAGTGTCTTACCAAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1171
 Db 1675 CAGAGAGATGGGAG 1734
 QY 1172 AATGAGATGAG 1231
 Db 1735 CTGTGGCATGAAACCATCTCTTTTGTGTCTTAATGACAGCTAAACAGCTATTGTGCTGCTGCT 1794

QY 1232 CCTGGCCATCTGTGATATTTGTGAGAGCACTTCCAGAGACCCATAGTTGCTGGCTG 1291
 Db 1795 TGTGTGGAG 1854
 QY 1292 GGGGCGCCAGTGTCTCTGTATGCTGACACTGCTGTGTGGAGCATATGATGATGATGATG 1351
 Db 1855 GGGGTAG 1914
 QY 1352 TGTAG 1411
 Db 1915 GCTTCTTGGGAG 1974
 QY 1412 GCTTCAATTTGATCTCTGATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1471
 Db 1975 GCTTGGCATCTTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2034
 QY 1472 AGATGGCAGCGCTTCCAGCGCGCTACAGCTACTGAGTGGTCCCTACAGAGGCGCTTACG 1531
 Db 2035 AGTTTCTACAG 2094
 QY 1532 TCTCTATGAGATTTATCAGAGTACTCAAAAGTGTATGATGATGATGATGATGATGATGATG 1591
 Db 2095 TATTCGATTTGTGAGTCTCAGAGTCTCAGAGTCTCAGAGTCTCAGAGTCTCAGAGTCTCAG 2154
 QY 1592 GGCCTCAGTCCACCCCGCTACTAG 1651
 Db 2155 GAGCCCGACAG 2214
 QY 1652 ACCCAACATACAG 1711
 Db 2215 CAGAGGTCAATCTGTAGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2274
 QY 1712 TGGGTTTCAAGTGTGCTTACATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1771
 Db 2275 GTACTTGTGTGAGAGTACTGAGTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2334
 QY 1772 TCATACCTTCAG 1831
 Db 2335 GTGTACCTTCAG 2394
 QY 1832 CTGCTTACAG 1891
 Db 2395 GCATTTACAG 2454
 QY 1892 TGGCATCAG 1951
 Db 2455 GCGGCGGCTTACAG 2514
 QY 1952 GCAACAGCGCGCGCTTCAACAGAGATGATCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2007
 Db 2515 TCACATATACAG 2570

RESULT 14
 US-09-819-946-3
 : Sequence 3, Application US/09819946
 : Patent No. US2002045740A1
 : GENERAL INFORMATION:
 : APPLICANT: Walke, D. Wade
 : APPLICANT: Scoville, John
 : APPLICANT: Donoho, Gregory
 : APPLICANT: Turner, C. Alexander Jr.
 : APPLICANT: Friedrich, Glenn
 : APPLICANT: Abulid, Alejandro
 : APPLICANT: zambrowicz, Brian
 : APPLICANT: Sands, Arthur T.
 : TITLE OF INVENTION: No. US2002045740A1 Human 7TM Proteins and Polynucleotides E
 : FILE REFERENCE: LEX-0157-USA
 : CURRENT APPLICATION NUMBER: US/09/819,946
 : CURRENT FILING DATE: 2001-08-27
 : PRIOR APPLICATION NUMBER: US 60/192,978
 : PRIOR FILING DATE: 2000-03-28

NUMBER OF SEQ ID NOS: 9
 SOFTWARE: FastSeq for Windows Version 4.0
 SEQ ID NO 3
 LENGTH: 2292
 TYPE: DNA
 ORGANISM: homo sapiens
 US-09-819-946-3

Query Match 19.2%; Score 386.4; DB 10; Length 2292;
 Best Local Similarity 51.1%; Pred. No. 2.7e-99;
 Matches 1042; Conservative 1; Mismatches 947; Indels 48; Gaps 4;

1 ATCACTACAGGCGCATGATGAGTGGAGACAGAGGTCGCTCCGCTTGGC 60
 265 ATTACTTGGGGCCAGCAGCAGACGCTCAGCGGAGGCGGACATATCCCTTCTG 324
 61 CGTACCAACCCAGCGGACGACACGATGAGGCTATGTCAGCTGATGCTGACATTC 120
 325 CGCACCATCCCAATGACAAATGACAGGTCAGGACCATGTCCTGCTGACAGAAATTC 384
 121 CGCTGAGACTGGATGATGCTGCTGAGCAGCGACCATATGCGCGGACAAATGGCCAG 180
 385 GGTGAGACTGATATCTCTGCTTGGCAGCTGACACTATGGCAGCTAGGGGTGCG 444
 181 CTGCTTGGCGAGCGCGTGGCCGCGCAGACATCTGATCGCTTCCAGAGAGCGTGGCC 240
 445 GCACCTGGAGAACCCAGGCGCAGTGTGAGGGGATCTGCTATTCATTCAGGACATCATATGCC 504
 241 ACACGACGCCCAACCAACATGACGTCAAGAGGCGCCAGCGCTGTGACCATTTG 300
 505 -----TTCTGTGCCAGGTGGCGATGAGAGATGACATGCTCATG 546
 301 GACAAGCTGACAGACAGACAGCGCGCTGTCGTGCTGCTTCCCGACCTGACCTG 360
 547 CGCCACCTGGGCCAGCGCGGCGCCACGTCGTGTTTTCAGCGGAGATTTGGCC 606
 361 TACCACTTCTCAATGAGGTGCTGGCCAGAACTTACAGGGCGCGGTGATTCGCTTC 420
 607 AGGGTCTTTTTCAGTCCGTGGTGTGACCAACCTGACCTGAGGAGTGGTGGCTGCTCA 666
 421 GAGTCTTGGGCGATGACCCGCTCTGCAACACTCAGGAGCTGGGCCATTTGGCCACC 480
 667 GAAAGCTGGGCGCTTCCAGGACATCACTGGGGTGGCCGGATCCAGGCAATTTGGAGT 726
 481 TTCTGGGCAATCACCATCAGAGCTGCGGCTCCGCGCTTATGATGATTCGGCAGTGG 540
 727 GTGCTGGGCGGTGCGCATTCAGAGAGAGGGGTGTCTGCTGCTTAAGCGCTTAAAGAAC 786
 541 GGGCCACAGGCTGGGCGCCGACCTCTCAGCAGGACGACGAGCTAT-----ACCTGC 594
 787 TATGCGCCGGGAGCAAGAGGCGCCCTAGGCTTGCACAAAGGCTCTGTGTGACAGAC 846
 595 AACCAAGAGTGGACAACTGCTGAAAGCGCACTTGTCTTACAAACCATTTTCAGGCTC 654
 847 AATCAGCTCTGAGAAATGCGCAAGCTTTCATGRCACACAGCATGCCCAAGTCAAGGCG 906
 655 TCTGGGGAGCGTGTGCTACAGCTGTACTGCTGCTATGCTATGCTGGCCATCCCTG 714
 907 TTCTCCATGAGTGTCTGCTACAAAGCATACCGGGCTGTGTATGCGGTGGCCCATGGCC 966
 715 CACAGCTCTCTGCTGTGACAAAAGCACCTGACCAAGAGGGGTGTACCCCTGGCAG 774
 967 CACCGCTCTCTGGGCTGCTGCTGCTGAGCTTGTCCAGGGCGCGGATTCATCCCTGGCAG 1026
 775 CTGCTTGGAGAGATGGAAGGTCAACTTCACTGCTGAGACCAACAATCTTCTGAC 834
 1027 CTTTGGAGACATCACAAGGTGATTTCTTTCACACAAAGACATGCTGGCGTTTAT 1086
 835 CGCAAGGAGAGTGGCTCTGCACTTGGAGATTTGCAATGGAGACGGAGCGAG 894
 1087 GACAACAGAGATCCCTCAGTAGTATAAATATTTGCTGGAGCTGGAGATGAGACCAAG 1146
 895 AATCCCTTCCAGAGGCTGCTCTAC-----TACCCCTGACGACAGCTGAAGAC 948

1147 TGGACCTTCAGGCTCCTGCTGCTCCTCAGACATGCTGCCAGTTACAGTAAACATATATG 1206
 949 ATCAAGACATCTCTGACACCGCTCAACAAACAGATCCCTATGCTTCCATGTTCCAGAG 1008
 1207 ACCAAATTCAGTGGCAGGAAAGCAACCAAGTCCATGCTGTGTTCCAGCGAC 1266
 1009 TGGCAGTCAAGGCAAAAGAAAGAGGCTGTGGGCAATCCAGCTGCTGCTTGAAGGCAATC 1068
 1267 TGTCTTGAAGGGCACCGAGGAGTGTACGGGTTCATCACTGCTGCTTGAAGTGTG 1326
 1069 GACTGCTTCCGGGACCTTCCCTCAACACACTGAA-----TGGCCG 1110
 1327 CCTGTGGGGGCTGGGACCTTCTCAACAAAGATGACCTTACAGATGCCAGGCTGTGG 1386
 1111 AATAAGAGTGTCTTACAGAGTGAACCTCTGCTTCAAGCGGAGCTGTCTCTG 1170
 1387 AAAGAAAGTGGGACCTGAGGAAAGCCAGACCTGCTCCGCACTGTGTGTTTGG 1446
 1171 GAATGGCATGAGGCAACCCACATGCTGTGGGCGCTGCTGGCCCTGGGCTTCCAGAC 1230
 1447 GCTTGGCTGACACACACTTGTGGGTGCTGTGCGACGCTAACAGCTGCTGCTGCTG 1506
 1231 ACCCTGGCCATCTGTGATATTTGAGGCACTTCCAGACACCCATGATGCTGCTG 1290
 1507 CTGCTTGGGAGTGTGGCTGTTTGGCTGACACTAGACACCCCTGTGTGAGATGACGA 1566
 1291 GGGGGCCCATATGCTCTCTGATGCTGACACCTGCTGCTGTGTGTGATACATGCTGCT 1350
 1567 GGGGGCCGCTTGTCTTCTTATGCTGTGGCTCTCCGACAGAGTGTGAGCCCTTAT 1626
 1351 GTGTAGTGGGGCCCGCCAGAGTCTCCACCTGCTCTGCGGCGAGGCGCTTCCCTC 1410
 1627 GCTTCTTGGGGAAACCCAAAGGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCT 1686
 1411 TGCCTCAATTTGATCTCTGTATGCGGCTGCTTCTTCAATGCTGCTGCTGCT 1470
 1687 GCTTACACATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1746
 1471 AAGATGGCGAGCGCTTCCACAGCGCTGACACTACTGGGTGCTGCTGCTGCTGCTGCT 1530
 1747 AAGTTTCCACCAAGGTACTCATATTCACAGCGCTGGGTCCAAACACGAGTGTGCT 1806
 1531 GTCTATGAGCAATTTATACAGGTACTCAAAATGATGATGTGATTTGGATGCTGCA 1590
 1807 CTGTTGTGATATGACGCTCAGCGGCGCGCTGCTTATGCTGCTGCTGCTGCTGCT 1866
 1591 CGGCTCAGTCCACCCCGTACTGACCCGATGACCCCAAGATCAAAATGCTCTGCT 1650
 1867 TGGACCCCACTGCTGCTGAGGAATACAGCGCTTCCCAATCTGTGATGCTGAGTGC 1926
 1651 AACCCCACTACCGAAGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1710
 1927 ACAGAGACCAACTCCCTGGGCTTCAATGAGCTTCTTCAAAAGGCTCTCTTCATC 1986
 1711 GTGGGTTTCAAGCTTGGCTTACATGGGCAAAAGATGCTGCCCAACATCAACAGAGGCAAG 1770
 1987 AGTGGCTTGGCTTGGAGTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2046
 1771 TTCAATCAGCTTACGATGATGCTTATTTACCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1830
 2047 TGTGTACCTTACGCTGCTGCTTCAACTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2106
 1831 TCTGCTACAGGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1890
 2107 ACCTTCTACAGAGGCAAGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2166
 1891 CTGGCATCAGCTGCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1950
 2167 ACAGAGGCTTGGGTGATTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2226
 1951 CGCAACGCGCGCTTACTTCAACAGATGATCAGGCTTACACATGAGAGGAGCT 2008

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 23, 2003, 11:21:44 ; Search time 6591.01 Seconds

(without alignments)
11299.347 Million cell updates/sec

Title: US-09-927-315-14

Perfect score: 2559

Sequence: 1 atgcgggcccgtctcctc.....atcagggaacatgagtgta 2559

Scoring table: IDENTITY_NUC

Gaped 10.0 , Gapped 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb_da:*

2: gb_hcg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sts:*

12: gb_sy:*

13: gb_un:*

14: gb_vl:*

15: em_da:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_om:*

21: em_or:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vl:*

30: em_hcg_hum:*

31: em_hcg_inv:*

32: em_hcg_other:*

33: em_hcg_mus:*

34: em_hcg_pln:*

35: em_hcg_rod:*

36: em_hcg_mam:*

37: em_hcg_vrt:*

38: em_sy:*

39: em_hlgo_hum:*

40: em_hlgo_mus:*

41: em_hlgo_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2557.4	99.9	3458	6	AX417279
2	2555.8	99.9	2559	6	AX239649
3	2555.8	99.9	2559	6	AX298370
4	2555.8	99.9	3489	6	AX239647
5	2514.4	98.3	2729	6	AX354026
6	2090.6	81.7	2739	6	AX318243
7	1949	76.2	4256	9	AB065647
8	1949	76.2	84445	9	AL139287
9	1947.4	76.1	8001	6	AX417281
10	1923.8	75.2	197748	2	AC026283
11	1574	61.5	1788	6	AX451560
12	1470.8	57.5	2577	10	AF456324
13	1470.8	57.5	2577	10	AY032620
14	1454.8	56.9	2577	6	AX298367
15	1454.8	56.9	2577	10	AF311386
16	1454.8	56.9	3521	10	AF337039
17	1454.8	56.9	3554	10	AB049994
18	1450	56.7	2577	10	AY032621
19	1083	42.3	1113	6	AX451557
20	873.2	34.1	3150	10	AY026318
21	871.6	34.1	3153	10	AF368024
22	871.6	34.1	4105	10	AB055708
23	871.6	34.1	11809	6	AX298368
24	871.6	34.1	232941	2	AL670236
25	871.6	34.1	254197	2	AF389853
26	870	34.0	3154	10	AF368025
27	621	24.3	190537	2	AC126156
28	366	14.3	2526	6	AX282892
29	366	14.3	2951	6	AX282900
30	365.6	14.3	2759	6	AX354027
31	359	14.0	2516	9	AF387618
32	357	14.0	2292	6	AX282894
33	354	13.8	2526	10	AF301162
34	350.8	13.7	2520	10	AF127389
35	346	13.5	2892	10	AF337040
36	344.4	13.5	2529	10	AY032622
37	341.2	13.3	2526	10	AF301161
38	328.4	12.8	190537	2	AC126156
39	317.2	12.4	98248	2	AC095793
40	296.8	11.6	3060	10	AF337041
41	284	11.1	2532	6	AR208616
42	284	11.1	2532	10	AY032623
43	279.4	10.9	2529	10	AF127390
44	279.4	10.9	2993	6	AR208615
45	218.6	8.5	2010	6	AR208617

ALIGNMENTS

RESULT 1

AX417279

LOCUS AX417279 3458 bp DNA linear PAT 18-JUN-2002

DEFINITION Sequence 1 from Patent WO0230981.

ACCESSION AX417279

VERSION AX417279.1 GI:21522589

KEYWORDS

SOURCE

ORGANISM human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1

DIFFERENCES

Gene encoding a human g-protein coupled receptor and its use

Patent: WO 0230981-A I 18-APR-2002;

FEATURES	PE CORP NY (US)	Location/Qualifiers
SOURCE	1..3458	/Organism="Homo sapiens" /db_xref="taxon:9606"
BASE COUNT	590 a 1179 c 1092 g 597 t	
ORIGIN		
Query Match	99.9%; Score 2557.4; DB 6; Length 3458;	
Best Local Similarity	100.0%; Pred. No. 0;	
Matches 2558; Conservative	Mismatches 0; Indels 1; Gaps 0;	
QY	1 ANCTGGGCGCCTCGTCTCTGGGCTCAGCCCTCTGGGCTCTCTGCACCCCTTGGACGGGG	60
DB		
DB	30 ATGCTGGGGCCCTGCTGTCTCTGGGCCCTCAACCCTCTGGGCTCTCTGCACCCCTGGAGGGG	89
QY	61 GCCCATTTGCTGTCAACAGCAACTTGAAGAAGGGGAACTAGTGCTGGGGGGCTG	120
DB		
QY	121 TTCGCCCTGGGGAGAGGCCAGAGAGCTGGCTCCACAGCGGAGACAAGGCCACGACCT	180
DB		
DB	150 TTCCCCCTGGGGAGAGGCCAGAGAGGCTGGCTCCCGAAGCCGAGACAGGCCCT	209
QY	181 GTGTGACCAAGTTCTCTCAAACGGCTGTCTTGGGCACTGGCCATGAATAATGGCCGTG	240
DB		
DB	210 GTGTGCACCAAGTTCTCTCAAACGGCCGTCTGGGCACTGGCCATGAATAATGGCCGTG	269
QY	241 GAGGAGATCAACAAGTCGATGCGTCCCGGGCTGGCGCTGGGCTAGACCTCTTT	300
DB		
DB	270 GAGGAGATCAACAAGTCGATGCGTCCCGGGCTGGCGCTGGGCTAGACCTCTTT	329
QY	301 GATPAGTGCTCGAGACCTGTGTGGCCATGAAGCCAGCCAGCTATGTCTCTGGCCAAGCA	360
DB		
DB	330 GATPAGTGCTCGAGACCTGTGTGGCCATGAAGCCAGCCAGCTATGTCTCTGGCCAAGCA	389
QY	361 GGCAGCCCGGACATGGCCCTCTACTGCAACTACAGCACTACAGCCCGCTGTGCTGCT	420
DB		
DB	390 GGCAGCCCGGACATGGCCCTCTACTGCAACTACAGCACTACAGCCCGCTGTGCTGCTG	449
QY	421 GTCATCGGGGCCCACGTCGTACAGGTCGGCCATGAGTACCGGGCAAGTTCACGTTCTTC	480
DB		
DB	450 GTCATCGGGGCCCACGTCGTACAGGTCGGCCATGAGTACCGGGCAAGTTCACGTTCTTC	509
QY	481 CTCATGCCCCAGGTCAGCTACGCTGTACATGAGAGCTGTCTAGGCGCCGGAGACCTTC	540
DB		
DB	510 CTCATGCCCCAGGTCAGCTACGCTGTACATGAGAGCTGTCTAGGCGCCGGAGACCTTC	569
QY	541 CCTCTCTTCTTCGCAACCGTCGCCAGCGAGCTGTGACGTGACGGCCGCCGGAGCTG	600
DB		
DB	570 CCTCTCTTCTTCGCAACCGTCGCCAGCGAGCTGTGACGTGACGGCCGCCGGAGCTG	629
QY	601 CTGCAAGAGTTCGGTGAACGTGGTGGCCCGCTGGGACAGGACAGAGTAGACGGCGG	660
DB		
DB	630 CTGCAAGAGTTCGGTGAACGTGGTGGCCCGCTGGGACAGGACAGAGTAGACGGCGG	689
QY	661 CAGGCGCTGAGCATTTCTCGGCGCTTGGCCGGAGCGAGCGGAGCATGTGCATCGCAGCAG	720
DB		
DB	690 CAGGCGCTGAGCATTTCTCGGCGCTTGGCCGGAGCGAGCGGAGCATGTGCATCGCAGCAG	749
QY	721 GGCTGTGTGCGCGTCCCGGTGCGATGATCTCGCGCTGGGGAAAGTGCAGGACGTCTG	780
DB		
DB	750 GGCTGTGTGCGCGTCCCGGTGCGATGATCTCGCGCTGGGGAAAGTGCAGGACGTCTG	809
QY	781 CACCAAGGTGAACCAAGAGAGCGTGAAGTGTGTGCTGTTCGCTCCGTGACGCGCC	840
DB		
DB	810 CACCAAGGTGAACCAAGAGAGCGTGAAGTGTGTGCTGTTCGCTCCGTGACGCGCC	869
QY	841 CACGCGCTCTTCAATTACAGATCAGACAGCAAGGCTCTGCGCCAAAGTGTGGGTGGCCAGC	900
DB		
DB	870 CACGCGCTCTTCAATTACAGATCAGACAGCAAGGCTCTGCGCCAAAGTGTGGGTGGCCAGC	929
QY	901 GAGGCTGTGGCTGACCTGTGACTGTGATGTGGGGCTGCCCGGAGTGGCCAGATGGGACAG	960

[illegible]

Db	2010	GAGTCACAAACTGCGCTCTGAGCTGGGACAGACGGGCTGAGTGCGCTCTGGGGGGCCCTGG	2069
QY	2041	GCCGTGAGTGTGTGTGTGCTGTGGCCATGCTGTGTGAAGTGTGCACGTGTGACCTGTGTACTTG	2100
Db	2070	GCGTGTGTGGTGGTGGTGTGCTGTGGCCATGTGCTGTGTGAAGTGTGCACGTGTGTGACCTGTGTACTTG	2129
QY	2101	GTGGCCCTTCCCGCCGGAGGTGTGTACGAGTGTGGACATGTCTGTCCACAGGAGGCGCTGTGTG	2160
Db	2130	GTGGCCCTTCCCGCCGGAGGTGTGTACGAGTGTGGACATGTCTGTCCACAGGAGGCGCTGTGTG	2189
QY	2161	CACGTCCGACACAGCTCTGTGTGGTGCACCTTCGGGCTAAGGCGACGCAACCAATGTGCCACGTG	2220
Db	2190	CACGTCCGACACAGCTCTGTGTGGTGCACCTTCGGGCTAAGGCGACGCAACCAATGTGCCACGTG	2249
QY	2221	GCCCTTCTGTGCTCTGTGTGGGCACTTTCCTGTGTGTGGAGACGACCGCGGTGTGTACAAACGT	2280
Db	2250	GCCCTTCTGTGCTCTGTGTGGGCACTTTCCTGTGTGTGGAGACGACCGCGGTGTGTACAAACGT	2309
QY	2281	GCCCGTGGCCCTCACCTTGTGCCATGTGAGCCTACCTTATCACCCTGGGTCTCCTTGTGTGCC	2340
Db	2310	GCCCGTGGCCCTCACCTTGTGCCATGTGAGCCTACCTTATCACCCTGGGTCTCCTTGTGTGCC	2369
QY	2341	CTCTGTGGCAATGTGAGAGTGTGTCTCAGGACCCGCGCTGTCAATGTGGGGCCCTCTGTCTC	2400
Db	2370	CTCTGTGGCAATGTGAGAGTGTGTCTCAGGACCCGCGCTGTCAATGTGGGGCCCTCTGTCTC	2429
QY	2401	TGTGTCTCGGGGACATCGAGGTGTGCCTTCACATGTGCCACAGGTGTACTGTCTCATGTGGGAG	2460
Db	2430	TGTGTCTCGGGGACATCGAGGTGTGCCTTCACATGTGCCACAGGTGTACTGTCTCATGTGGGAG	2489
QY	2461	CCAGGGGCTCAACACCCCGAGTCTTCTGTGTGAGGAGGGGCGCTGTGGGATGTGCCAAGGCGAG	2520
Db	2490	CCAGGGGCTCAACACCCCGAGTCTTCTGTGTGAGGAGGGGCGCTGTGGGATGTGCCAAGGCGAG	2549
QY	2521	AATGACGGGAACACAGGAATTCAGGGGAAACATGAGTGA	2559
Db	2550	AATGACGGGAACACAGGAATTCAGGGGAAACATGAGTGA	2588

RESULT 2	AX239649	AX239649	2559 bp	DNA	linear	PAT 26-SEP-2001
LOCUS	AX239649					
DEFINITION	Sequence 15 from Patent.		W00164882.			
ACCESSION	AX239649					
VERSION	AX239649.1	GI:15797322				

SOURCE ORGANISM	human.
HOMO S	

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE	AUTHORS	TITLE
1 (bases 1 to 2559)	Glücksman, M.A., Galvin, K.M. and Silos-Santiago, I.	1993, 52881, 2398, 45449, 50289, and 52872, G protein-coupled

JOURNAL
Patent: WO 0164882-A 15 07-SEP-2001;
Millennium Pharmaceuticals, Inc. (US)

FEATURES	Location/Qualifiers
source	1. .2559

BASE COUNT	381	a	863	c	840	g	47
ORIGIN	/organism="Homo sapiens" /db_xref="taxon:9606"						

Query Match	99.98	Score 2555.8	DB 6	Length 2559
Best Local Similarity	99.98	Pred. No. 0		
Matches 2557	Conservative	0	Mismatches 2	Indels 0
				Gaps 0

0y 1 ATCTGAGGACCTGCTGTCCTTGGGCTCAAGCTCTGGGCTCTCTCTGACACCTTGAGACGGG 60

Db 1 ATGCTGGGACCTGCTGTCCTTGGGCTCAAGCTCTGGGCTCTCTCTGACACCTTGAGACGGG 60

0y 61 GCCCATGTGCTGCTGCTACACCAACTTAGATGAAGGGGACTTCGTCTGGGGGGCTG 120

Db	61	GC	CC	CA	TT	GT	TC	CT	CT	GA	CA	CA	CTT	AG	AT	GA	GA	GG	GA	CT	AC	TG	CT	G	GG	GG	GG	CT	G	120	
Qy	121	TT	CC	CC	CC	GG	GG	GG	GG	GG	GG	GG	GG	CT	CC	GA	GG	CC	GA	CA	GG	CC	CA	GG	CC	CA	GG	CC	CA	180	
Db	121	TT	CC	CC	CC	GG	GG	GG	GG	GG	GG	GG	GG	CT	CC	GA	GG	CC	GA	CA	GG	CC	CA	GG	CC	CA	GG	CC	CA	180	
Qy	181	GT	GT	CA	CA	CA	GG	TT	CT	CT	CA	AA	GG	CT	CT	GT	GG	GA	CA	CT	GG	CC	CA	TG	AA	AA	TG	CC	GT	240	
Db	181	GT	GT	CA	CA	CA	GG	TT	CT	CT	CA	AA	GG	CT	CT	GT	GG	GA	CA	CT	GG	CC	CA	TG	AA	AA	TG	CC	GT	240	
Qy	241	G	A	G	G	A	G	A	T	C	A	A	G	T	G	G	A	T	C	T	C	C	C	C	C	C	C	C	C	C	300
Db	241	G	A	G	G	A	T	C	A	A	G	T	G	G	A	T	C	T	C	C	C	C	C	C	C	C	C	C	C	C	300
Qy	301	G	A	T	A	G	T	C	T	G	A	G	A	C	T	T	G	T	G	C	C	A	T	G	T	C	T	G	C	C	360
Db	301	G	A	T	A	G	T	C	T	G	A	G	A	C	T	T	G	T	G	C	C	A	T	G	T	C	T	G	C	C	360
Qy	361	G	G	C	A	C	C	C	G	C	A	T	C	T	G	C	C	A	T	C	T	G	C	C	A	T	C	T	G	C	420
Db	361	G	G	C	A	C	C	C	G	C	A	T	C	T	G	C	C	A	T	C	T	G	C	C	A	T	C	T	G	C	420
Qy	421	G	T	C	A	T	C	G	G	G	G	C	C	A	T	C	T	G	C	A	G	A	C	T	T	G	C	T	T	C	480
Db	421	G	T	C	A	T	C	G	G	G	C	C	A	T	C	T	G	C	A	G	A	C	T	T	G	C	T	T	C	480	
Qy	481	C	T	C	A	T	C	G	C	C	A	G	T	C	A	G	T	G	C	A	T	G	A	G	A	C	T	T	C	540	
Db	481	C	T	C	A	T	C	G	C	C	A	G	T	C	A	G	T	G	C	A	T	G	A	G	A	C	T	T	C	540	
Qy	541	C	C	T	C	T	T	T	C	C	G	A	C	C	G	T	G	C	A	G	T	G	C	A	G	C	C	G	C	600	
Db	541	C	C	T	C	T	T	T	C	C	G	A	C	C	G	T	G	C	A	G	T	G	C	A	G	C	C	G	C	600	
Qy	601	C	T	G	C	A	G	A	T	T	G	G	G	T	G	G	G	C	C	T	G	G	G	A	G	A	C	A	G	660	
Db	601	C	T	G	C	A	G	A	T	T	G	G	G	T	G	G	G	C	C	T	G	G	G	A	G	A	C	A	G	660	
Qy	661	C	A	G	G	C	T	G	A	C	A	T	T	T	C	G	G	C	T	G	C	G	C	G	G	A	C	A	T	720	
Db	661	C	A	G	G	C	T	G	A	C	A	T	T	T	C	G	G	C	T	G	C	G	C	G	G	A	C	A	T	720	
Qy	721	G	G	C	T	G	T	G	C	C	G	C	C	G	A	T	G	A	T	C	C	G	G	C	T	G	G	G	A	780	
Db	721	G	G	C	T	G	T	G	C	C	G	C	C	G	A	T	G	A	T	C	C	G	G	C	T	G	G	G	A	780	
Qy	781	C	A	C	A	G	G	A	C	A	G	A	G	G	T	G															

QY	1201	GCAGAGCCCTGCACAACTCTTCAGTGCACACGCTCAGGCTGCGCCGCGAGACCCC	1280
Db	1201	GCCCCAGCCCTGCACAACTCTTCAGTGCACACGCTCAGGCTGCGCCGCGAGACCCC	1260
QY	1261	GTCAGACCCCTGCAGGCTCTGAGAAACATGTCAACCTACCTTCACGTGGGCGGGCTG	1320
Db	1261	GTCAGACCCCTGCAGGCTCTGAGAAACATGTCAACCTACCTTCACGTGGGCGGGCTG	1320
QY	1321	CCGCTGCGGTTCCAGACAGCGGAAACGTGACATGACCTTGACGTGTGGGTG	1380
Db	1321	CCGCTGCGGTTCCAGACAGCGGAAACGTGACATGACCTTGACGTGTGGGTG	1380
QY	1381	TGCGAGGAGCTCAGGCCCCAGGCTCCAGACAGTGGGAGGTTCAAGGGCAGCCTCAGAGCA	1440
Db	1381	TGCGAGGAGCTCAGGCCCCAGGCTCCAGACAGTGGGAGGTTCAAGGGCAGCCTCAGAGCA	1440
QY	1441	GAGCGCCTGAAGATCCGCTGCGCACAGCTGTGACCAACAGAAAGCCGCTGCCGGTGTCTG	1500
Db	1441	GAGCGCCTGAAGATCCGCTGCGCACAGCTGTGACCAACAGAAAGCCGCTGCCGGTGTCTG	1500
QY	1501	CGGAGGTGCAGAGGGGCGCAGGTGGCGCCGGGTCCAAAGGGTTCCACTCCGCTGCTAGCAC	1560
Db	1501	CGGAGGTGCAGAGGGGCGCAGGTGGCGCCGGGTCCAAAGGGTTCCACTCCGCTGCTAGCAC	1560
QY	1561	TGTGTGACATTCGAGAGGGGCGACCTACCGGCAAAACCCAGACATCGCTGCACCTTT	1620
Db	1561	TGTGTGACATTCGAGAGGGGCGACCTACCGGCAAAACCCAGACATCGCTGCACCTTT	1620
QY	1621	TGTGGCCAGGATGATGTGTTCCCGGAGCGCAACACGCTCTCCGCGCGACAGTCTCGG	1680
Db	1621	TGTGGCCAGGATGATGTGTTCCCGGAGCGCAACACGCTCTCCGCGCGACAGTCTCGG	1680
QY	1681	TTCTGTCGATAGGAGCGGAGCGGCTGTAGTCTCTCTCTCTGACTGACCTGGCGGTG	1740
Db	1681	TTCTGTCGATAGGAGCGGAGCGGCTGTAGTCTCTCTCTCTCTGACTGACCTGGCGGTG	1740
QY	1741	GGCCTTGTGCTGCGCTGCTTTTGGGGCTGTTCGTTACACATCGGAGACGCCACTGGTTCAG	1800
Db	1741	GGCCTTGTGCTGCGCTGCTTTTGGGGCTGTTCGTTACACATCGGAGACGCCACTGGTTCAG	1800
QY	1801	GCTCGGGGGGGGGCCCGGGGCTGCTTTTGGGCTGTGTGGCTGGGGGCTGGTCTGCCTCAGC	1860
Db	1801	GCTCGGGGGGGGGCCCGGGGCTGCTTTTGGGCTGTGTGGCTGGGGGCTGGTCTGCCTCAGC	1860
QY	1861	GTCCTCTGTTCCTCTGGGCGACGCCACGCCCTGCGCATGCTGGCCAGACGCCCTTGTCC	1920
Db	1861	GTCCTCTGTTCCTCTGGGCGACGCCACGCCCTGCGCATGCTGGCCAGACGCCCTTGTCC	1920
QY	1921	CACCTCCCGCTCAACGGGCTGCTGAGCAACAATCTTCGTGCAAGCGCGCGAAGATCTTCGTG	1980
Db	1921	CACCTCCCGCTCAACGGGCTGCTGAGCAACAATCTTCGTGCAAGCGCGCGAAGATCTTCGTG	1980
QY	1981	GAGTCAGAACTGCTCTGAGACTGGGCGACCGGCTGATGTGGTGGCTGGGGGGGCGCTCGG	2040
Db	1981	GAGTCAGAACTGCTCTGAGACTGGGCGACCGGCTGATGTGGTGGCTGGGGGGGCGCTCGG	2040
QY	2041	GCCCTGCGTGTGTGTGTGCTGTGCGGCAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2100
Db	2041	GCCCTGCGTGTGTGTGTGCTGTGCGGCAATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2100
QY	2101	GTCGACCTTCGCGGAGGT	2160
Db	2101	GTCGACCTTCGCGGAGGT	2160
QY	2161	CACGTGCGACACGCTCTGTGGGTGACGCTTCGCGCTTACGCGACGCCACCAATGTCCACGCTG	2220
Db	2161	CACGTGCGACACGCTCTGTGGGTGACGCTTCGCGCTTACGCGACGCCACCAATGTCCACGCTG	2220
QY	2221	GCCCTTCTGTGCTTCTCTGGGCACTTTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2280
Db	2221	GCCCTTCTGTGCTTCTCTGGGCACTTTCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2280

[illegible]

D	b	301	GATACGTCTCGAGAGCTGTGTGTGGCCATGAAAGCCAGCCTATGTTCCGTGGCCAAAGCA	360
O	y	361	GGCAGCCGCGACATCGCCGCGCTACTGCACATACACGACATACAGCCCGCGTGTGCTGACT	420
D	b	361	GGCAGCCGCGACATCGCCGCGCTACTGCACATACAGCAGTACAGCCCGCGTGTGCTGACT	420
O	y	421	GTCATCGGGGCCCCCATCTGTGAGAGCTCGCCATGTGTACCGGCAAGTTCTTACGTTCTTC	480
D	b	421	GTCATCGGGGCCCCCATCTGTGAGAGCTCGCCATGTGTACCGGCAAGTTCTTACGTTCTTC	480
O	y	481	CTCATGCCCCCAGGTCAGACTCAGGGTGTAGCATGGAGCTCTCAGCGCCCGGAGACCTTC	540
D	b	481	CTCATGCCCCCAGGTCAGACTCAGGGTGTAGCATGGAGCTCTCAGCGCCCGGAGACCTTC	540
O	y	541	CCCTCCTTCTTCCCGCACCAGCGGCCCCAGACCGGTGTGCACACTACGCGCGCCGAGACTG	600
D	b	541	CCCTCCTTCTTCCCGCACCAGCGGCCCCAGACCGGTGTGCACACTACGCGCGCCGAGACTG	600
O	y	601	CTGCAGAGAGTTCCGCTGGAACCTGGGTGGCCGCGCCTGGGAGGAGAGACAGAGCGCGG	660
D	b	601	CTGCAGAGAGTTCCGCTGGAACCTGGGTGGCCGCGCCTGGGAGGAGAGACAGAGCGCGG	660
O	y	661	CAGGGCCTGAGCATCTTCTCGGCCCTTGGCCCGCGGACGCGGCACTGTGATCGCGCAGAG	720
D	b	661	CAGGGCCTGAGCATCTTCTCGGCCCTTGGCCCGCGGACGCGGCACTGTGATCGCGCAGAG	720
O	y	721	GGCCTGGTGGCGCGTCCCGCGTGGCCATGACTCGCGCTGGGGAAGTGCAGAGACTCTG	780
D	b	721	GGCCTGGTGGCGCGTCCCGCGTGGCCATGACTCGCGCTGGGGAAGTGCAGAGACTCTG	780
O	y	781	CACCAAGGTGAACAGAGCAGCAGCTGTGAGAGTGTGCTCTTTCGCTCGCTGACAGCCGCC	840
D	b	781	CACCAAGGTGAACAGAGCAGCAGCTGTGAGAGTGTGCTCTTTCGCTCGCTGACAGCCGCC	840
O	y	841	CAGCCCTCTTCAACTACACATACAGACAGGACTCTCGCCCAAGGTGTGGGTGGCCAGC	900
D	b	841	CAGCCCTCTTCAACTACACATACAGACAGGACTCTCGCCCAAGGTGTGGGTGGCCAGC	900
O	y	901	GAGGCTCTGGTGACCTCTGACCTGTGATGTGGGGCTCCCGGATGGCCAGATGGGAGCAG	960
D	b	901	GAGGCTCTGGTGACCTCTGACCTGTGATGTGGGGCTCCCGGATGGCCAGATGGGAGCAG	960
O	y	961	GTCCTTGGCTTCTCTCAGAGAGGGTGCCAGCTGCACAGATTTCCCAAGTACGTGAAGC	1020
D	b	961	GTCCTTGGCTTCTCTCAGAGAGGGTGCCAGCTGCACAGATTTCCCAAGTACGTGAAGC	1020
O	y	1021	CACCTGGGCGCGGCGCACAGACCCGGGCTTGTGCTGGCCCTGGGGAGAGGGAGCAGGCT	1080
D	b	1021	CACCTGGGCGCGGCGCACAGACCCGGGCTTGTGCTGGCCCTGGGGAGAGGGAGCAGGCT	1080
O	y	1081	CTGAGAGAGGACGTGTGTGGGCGACGCGCTCGCGCAGATGTGATGTGATCAGCTGCAGAAC	1140
D	b	1081	CTGAGAGAGGACGTGTGTGGGCGACGCGCTCGCGCAGATGTGATGTGATCAGCTGCAGAAC	1140
O	y	1141	GTCAGCCAGAGGGCTAAATTCACCAACAGAGCTTCTGTGTACGACAGCTGTGTATAGCTG	1200
D	b	1141	GTCAGCCAGAGGGCTAAATTCACCAACAGAGCTTCTGTGTATGACAGCTGTGTATAGCTG	1200
O	y	1201	GCCAGGACCGCTGCACAACTTTTCAAGTCAACAGCCTCAGCTGCGCCGCGAGAGACCC	1260
D	b	1201	GCCAGGACCGCTGCACAACTTTTCAAGTCAACAGCCTCAGCTGCGCCGCGAGAGACCC	1260
O	y	1261	GTCAGAGCCCTTGACAGCTCTTGGAAGAACTGTACACCTTGACCTTCAAGTGTGGGCTG	1320
D	b	1261	GTCAGAGCCCTTGACAGCTCTTGGAAGAACTGTACACCTTGACCTTCAAGTGTGGGCTG	1320
O	y	1321	CCGCTGGGTTTCAGACAGCAGCCGGAAGCTGTGACATGAGATACAGCTGAGCTGTGGGTG	1380
D	b	1321	CCGCTGGGTTTCAGACAGCAGCCGGAAGCTGTGACATGAGATACAGCTGAGCTGTGGGTG	1380
O	y	1381	TGGCAGGAGCTCAGTGTGCCAGGCTCCACAGAGTGGGCAAGTTCAACGGAGCCTCAGACA	1440
D	b	1381	TGGCAGGAGCTCAGTGTGCCAGGCTCCACAGAGTGGGCAAGTTCAACGGAGCCTCAGACA	1440

[illegible]

QY 2521 AATGACGGGAAACAGGAAATCAGGGGAAACATGATGA 2559
Db 2521 AATGACGGGAAACAGGAAATCAGGGGAAACATGATGA 2559

RESULT 4
LOCUS AX239647 3489 bp DNA linear PAT 26-SEP-2001
DEFINITION Sequence 13 from Patent WO0164882.
ACCESSION AX239647
VERSION AX239647.1 GI:15797320
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
AUTHORS Glucksman, M.A., Galvin, K.M. and Silos-Santlago, I.
TITLE 1 (bases 1 to 3489)
JOURNAL receptors and uses thereof
PATENT: WO 0164882-A 13 07-SEP-2001;
Millemmium Pharmaceuticals, Inc. (US)
FEATURES
source location/Qualifiers
1..3489
/organism="Homo sapiens"
/db_xref="taxon:9606"
52..2610
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC8308.1"
/db_xref="GI:15797321"
/translation="MLGPAYLIGLSLNLHPGAPLICLSQJRMKGDVYLGILPLG
EAEAGLRSTRPSPVCTRFSSNGLLMALAMAVEINNKDILLGLGLDFT
CSBPVAMKSLMFLAKAGSRDIAACNYOXPRLVAVIGPSSLEAMTGFSEF
LMQVSGASMEILSAREPSEFRVSPYDOLTAALLOEFGVMVVALDSEY
GROGLSTESLAAARGICIAHEGLVPLPRADSDRIGVDOLHOVNOSSVOVILFAS
VHAHALFNYSISRLSPKWAPEAMTSDIYMGIPGAMOMTVIGTLDROQLHEF
POYKTHIALATPAPCSALGEROGIEEDVQRCPOCCITLONVSAGLNHTES
VYAAVSVAAQALNTLQCNASGCPADVPVWOLLEMYMLTFHVGLPLRFDSSGV
DMEYDLKLVWQSVPLRHVDVNGSLRTERLIRMTSDNKPVSRCRQCEGV
RYKFGHSCCYDCVDEAGSYRONPDIACTFCGDEWSPERSTRCPRRSRFLAME
PAYLLILLISLALGILVIALGLFVHHRDPLVOASGGLACGIVGLICVSYLLF
PGPSPARCLAOPLSHPLTGLSTFLFOALITPESELPISWADRLSCINGPMAN
LVYLLAMLVYALCTMTLVAPPEVYTDMLMLETALVHCRITSMVSEGLAHINATL
ALFCFTGLVRSQPGKYNRANRGLTFEAMLAFTFTWSFVLLANVOVLNPAVOMAL
LILCVLIGLAFHPLPRCYLIMROGLNTPERFLGSDADGODNGTNGKHE"

BASE COUNT 578 a 1189 c 1102 g 612 t 8 others
ORIGIN

Query Match 99.98; Score 2555.8; DB 6; Length 3489;
Best Local Similarity 99.98; Pred. No. 0;
Matches 2557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AATGTCGGCCCTGCTGCTCGGCTCAGCCTCTGGGCTCCTGCAACCTGGAGCGGG 60
Db 52 AATGTCGGCCCTGCTGCTCGGCTCAGCCTCTGGGCTCCTGCTGCAACCTGGAGCGGG 111
QY 61 GCCCATTTGCTGTGTACAGCACTTGAATGAAGGGGGAATAGTCTGGGGGGCTG 120
Db 112 GCCCATTTGCTGTGTACAGCACTTGAATGAAGGGGGAATAGTCTGGGGGGCTG 171
QY 121 TTCCCTTGGGCGAAGCGGAGGCTGGCTCCGACAGCGGAGACAGGCCCTT 180
Db 172 TTCCCTTGGGCGAAGCGGAGGCTGGCTCCGACAGCGGAGACAGGCCCTT 231
QY 181 GTGTGACACAGGCTCTCTCAAAAGGCTGTGTGGGACATGGGCAAGAAATAGCCGTG 240
Db 232 GTGTGACACAGGCTCTCTCAAAAGGCTGTGTGGGACATGGGCAAGAAATAGCCGTG 291
QY 241 GAGGAGATCAACAAGATGCTGTGCTCCGCGGCTGGCGCTGAGGCTAGACCTCTT 300
Db 292 GAGGAGATCAACAAGATGCTGTGCTCCGCGGCTGGCGCTGAGGCTAGACCTCTT 351

QY 301 GATAGTGCTGGAGACCTGTGTGTGGCCATGATGAGCCAGCCATGATTTCTTGCCAGAGCA 360
Db 352 GATAGTGCTGGAGACCTGTGTGTGGCCATGATGAGCCAGCCATGATTTCTTGCCAGAGCA 411
QY 361 GGCAGCGCGGACATCGCGGCTACTGCACTACAGAGTACAGAGCCCGGTGCTGGCT 420
Db 412 GGCAGCGCGGACATCGCGGCTACTGCACTACAGAGTACAGAGCCCGGTGCTGGCT 471
QY 421 GTCATCGGGGCCCCACCTGCTAGAGCTGCGCCATGATGTCACCGGCAAGTCTTACGCTTTC 480
Db 472 GTCATCGGGGCCCCACCTGCTAGAGCTGCGCCATGATGTCACCGGCAAGTCTTACGCTTTC 531
QY 481 CTCATGCCCAAGGTACGCTACGCTACGCTGCTACGATGAGTACGAGCTGCTGAGCGCCGAGACCTTC 540
Db 532 CTCATGCCCAAGGTACGCTACGCTACGCTGCTACGATGAGTACGAGCTGCTGAGCGCCGAGACCTTC 591
QY 541 CCCCTCTTCTTCCGACCGGCGCCAGGACCGGTGTGAGTGAACGGCGCGCGGAGACTG 600
Db 592 CCCCTCTTCTTCCGACCGGCGCCAGGACCGGTGTGAGTGAACGGCGCGCGGAGACTG 651
QY 601 CTGCAGAGATTCGCTGGAACCTGGGTGGCCGCTGGGCGAGCGAGCGAGTACGGCCGG 660
Db 652 CTGCAGAGATTCGCTGGAACCTGGGTGGCCGCTGGGCGAGCGAGCGAGTACGGCCGG 711
QY 661 CAGGCGCTGAGCATCTTCTGCGCCCTGCGCGGCGGCGAGCGGCAATCTGCGGCAAGAG 720
Db 712 CAGGCGCTGAGCATCTTCTGCGCCCTGCGCGGCGGCGAGCGGCAATCTGCGGCAAGAG 771
QY 721 GGCTGTGTCGCGCTGCGCCGCGGATGACCTGCGGGGTGGGGAAGGAGAGAGTCTCTG 780
Db 772 GGCTGTGTCGCGCTGCGCCGCGGATGACCTGCGGGGTGGGGAAGGAGAGAGTCTCTG 831
QY 781 CACCAAGTGAACAGAGACAGCTGACAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 832 CACCAAGTGAACAGAGACAGCTGACAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 891
QY 841 CAGCCCTCTTCAACTACGATACGATACGAGAGAGTCTGCGCCCAAGAGTGTGGTGGCCAGC 900
Db 892 CAGCCCTCTTCAACTACGATACGATACGAGAGAGTCTGCGCCCAAGAGTGTGGTGGCCAGC 951
QY 901 GAGGCTGCTGACCTCTGACCTGATGATGAGGCTGCGCCGATGAGGCGAGATGAGGCGAGC 960
Db 952 GAGGCTGCTGACCTCTGACCTGATGATGAGGCTGCGCCGATGAGGCGAGATGAGGCGAGC 1011
QY 961 GTGCTGCTTCTTCCACAGAGGGGTGCCACCTGACAGAGTTCCTCCAGTACGTGAAGAGC 1020
Db 1012 GTGCTGCTTCTTCCACAGAGGGGTGCCACCTGACAGAGTTCCTCCAGTACGTGAAGAGC 1071
QY 1021 CACCTGCGCTTGGCGACCGACCGGCGCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1072 CACCTGCGCTTGGCGACCGGCGCTTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1131
QY 1081 CTGAGAGAGAGAGTGTGTGGGCGACGCTGCGCGAGTGAATGATACATACAGCTTCAGAAAC 1140
Db 1132 CTGAGAGAGAGAGTGTGTGGGCGACGCTGCGCGAGTGAATGATACATACAGCTTCAGAAAC 1191
QY 1141 GTGAGCGCAGGGCTAAATCAACACACAGAGCTTCTGTCTACGAGCTGTGTATAGCGTGT 1200
Db 1192 GTGAGCGCAGGGCTAAATCAACACACAGAGCTTCTGTCTACGAGCTGTGTATAGCGTGT 1251
QY 1201 GCCAGGCGCTGCAACAACATCTTCACTGCAAGGCTGCAAGGCTGCGCGCGAGAGACCC 1260
Db 1252 GCCAGGCGCTGCAACAACATCTTCACTGCAAGGCTGCAAGGCTGCGCGCGAGAGACCC 1311
QY 1261 GTGAGCGCTGCGAGCTGCGAGGAGAAATGATGACAACTGACCTTCCAGTGGGGGGGCTG 1320
Db 1312 GTGAGCGCTGCGAGCTGCGAGGAGAAATGATGACAACTGACCTTCCAGTGGGGGGGCTG 1371
QY 1321 CCGCTGCGGTTTCAGACAGACGCGGAAAGTGTGAGATGAGTACGACCTGAAGCTGTGGGTG 1380
Db 1372 CCGCTGCGGTTTCAGACAGACGCGGAAAGTGTGAGATGAGTACGACCTGAAGCTGTGGGTG 1431
QY 1381 TGGCAGGGCTCAGTGTGCCAGGCTCCACAGCTGCGAGCGTGGGAGGTTCAACGGGACGCTCAGGACA 1440

Db	1432	TGGCAGGGCTCAGTGGCCAGAGCTCCAGACGATGAGGAGGTTTAAAGGCAAGCTTCAGAGCA	1491
Qy	1441	GAGCGCTGAAGATCCGATGGGCACACGTCTGACAAACAGAGCCGCTGTCGGGTCTCG	1500
Db	1492	GAGGCGCTGAAGATCCGCTGGCACACGTTGTGAAACAAACAAAGACCCTGTGCCGATGCTCG	1551
Qy	1501	CGGAGAGGCCAGAGGGCCAGAGGGCCGCGGATCAAGGGGTTCCACTCCGCTGCTGACAC	1560
Db	1552	CGGAGAGGCCAGAGGGCCAGAGGGCCGCGGATCAAGGGGTTCCACTCCGCTGCTGACAC	1611
Qy	1561	TGTGTGACTGCGAGGGGGGACACTACCGGCAAAACCCAGAGCATGCGCTGCACCTTT	1620
Db	1612	TGTGTGAGCTGCGAGGGGGGACACTACCGGCAAAACCCAGAGCATGCGCTGCACCTTT	1671
Qy	1621	TGTGGCCAGATGATGTGTCCCGGAGCGAAGACACAGCTGCTTCCGCGCAGGTCTCG	1680
Db	1672	TGTGGCCAGATGATGTGTCCCGGAGCGAAGACACAGCTGCTTCCGCGCAGGTCTCG	1731
Qy	1681	TTTCCTGGCATGGGGGCGAGCGGCTGTAGAGTGCAGTCTCTCTGAGCGTGGCGGTG	1740
Db	1732	TTTCCTGGCATGGGGGCGAGCGGCTGTAGAGTGCAGTCTCTCTGAGCGTGGCGGTG	1791
Qy	1741	GGCCTTGCTGTGGCTGCTTTGGGGCTTTCTTCAACATCGGAGACGCCACTGTGTTAG	1800
Db	1792	GGCCTTGCTGTGGCTGCTTTGGGGCTTTCTTCAACATCGGAGACGCCACTGTGTTAG	1851
Qy	1801	GCCCTGGGGGGGGCCCTGGGCTCTTTGGGCTGTGTGCTGTGGCTTGGTCTGCTCAAGC	1860
Db	1852	GCCCTGGGGGGGGCCCTGGGCTCTTTGGGCTGTGTGCTGTGGCTTGGTCTGCTCAAGC	1911
Qy	1861	GTCCTCTGTTCCTGTGGCCAGCCAGCCAGCCCTGCCAGTGCCTGAGAGCCCTGTGAC	1920
Db	1912	GTCCTCTGTTCCTGTGGCCAGCCAGCCAGCCCTGCCAGTGCCTGAGAGCCCTGTGAC	1971
Qy	1921	CACCTCCCGGCTCAAGGGCTGCTGAGCTGGGCGAGACACTTTCCTGACAGCGCGGAGTCTTG	1980
Db	1972	CACCTCCCGGCTCAAGGGCTGCTGAGCTGGGCGAGACACTTTCCTGACAGCGCGGAGTCTTG	2031
Qy	1981	GAGTCAGACATGCTCTGAGCTGGGCGAGACCGGCTGAGTGGCTGCTGCGGGGGCCCTGG	2040
Db	2032	GAGTCAGACATGCTCTGAGCTGGGCGAGACCGGCTGAGTGGCTGCTGCGGGGGCCCTGG	2091
Qy	2041	GCTGTGGTGTGTGTGCTGTGGCCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2100
Db	2092	GCTGTGGTGTGTGTGCTGTGGCCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT	2151
Qy	2101	GTTGGCTCTCCCGCGGAGGT	2160
Db	2152	GTTGGCTCTCCCGCGGAGGT	2211
Qy	2161	CACGTCCGACACGCTCTGTGGGTACGCTTCGCGCTAGCCAGCCACCAATGCCAGCGTG	2220
Db	2212	CACGTCCGACACGCTCTGTGGGTACGCTTCGCGCTAGCCAGCCACCAATGCCAGCGTG	2271
Qy	2221	GCTTCTGT	2280
Db	2272	GCTTCTGT	2331
Qy	2281	GCCCGTGGCCCTACACTTTGGCATGTGGGCTACTTATACATCGGTGCTCTTTGTGGC	2340
Db	2332	GCCCGTGGCCCTACACTTTGGCATGTGGGCTACTTATACATCGGTGCTCTTTGTGGC	2391
Qy	2341	CTCTCGGCATGT	2400
Db	2392	CTCTCGGCATGT	2451
Qy	2401	TGTGTCTGTGGGATCGGT	2460
Db	2452	TGTGTCTGTGGGATCGGT	2511
Qy	2461	CGAGGGCTCAACACCCCGAGTCTTCTGTGGAGGGGGCCGTGGGATGCCCAAGCGAC	2520

Db	2512	CCAGGGCTCAACACCCCGAGTTCTTCTCGGAGGGGGCCCTGGGGATGCCAAGGCCAG	2571
QY	2521	AATGACGGGACACACAGAAATCAGGGGAAATCATGTGTA	2559
Db	2572	AATGACGGGACACAGAAATCAGGGGAAATCATGTGTA	2610
RESULT 5			
LOCUS	AX354026	2729 bp	DNA
DEFINITION	Sequence 12 from Patent WO0203848.	Linear	PAT 06-FEB-2002
ACCESSION	AX354026		
VERSION	AX354026.1	GI:18618963	
KEYWORDS			
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Naito, A.T.		
AUTHORS	Method for opening the blood-brain barrier		
JOURNAL	Patent: WO 0203848-A 12 17-JAN-2002;		
FEATURES	Naito, Albert T. (US)		
source	Location/Qualifiers		
	1..2729		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/note="Incyte ID No: 7475194CB1"		
BASE COUNT	413 a 934 c 879 g 503 t		
ORIGIN			
Query Match	98.3%; Score 2514.4; DB 6; Length 2729;		
Best Local Similarity	98.7%; Pred. No. 0;		
Matches 2558; Conservative	0; Mismatches 1; Indels 33; Gaps 1;		
QY	1	ATGCTGCGCCCTGCTGCTCTGCGCCCTCAGCCTCTGGCCTCTCTGTCACCCCTGGACGGGG	60
Db	1	ATGCTGCGCCCTGCTGCTCTGCGCCCTCAGCCTCTGGCCTCTCTGTCACCCCTGGACGGGG	60
QY	61	GCCCATTTTGCTGCTGTCACAGCACAATTAGAGTGAAGGGGAGTACGTGCTGGGGGGGCTG	120
Db	61	GCCCATTTTGCTGCTGTCACAGCACAATTAGAGTGAAGGGGAGTACGTGCTGGGGGGGCTG	120
QY	121	TTTCCCTGGGGCGAGGCCGAGGAGGCTGGCTCCGACGCGGACAGCGGCCAGCAGCCCT	180
Db	121	TTTCCCTGGGGCGAGGCCGAGGAGGCTGGCTCCGACGCGGACAGCGGCCAGCAGCCCT	180
QY	181	GTGTGACACAGGTTTCTCTCAAAAGGCGCTGCTGGGCACTGAGCCATGAAAAATGGCCGCTG	240
Db	181	GTGTGACACAGGTTTCTCTCAAAAGGCGCTGCTGGGCACTGAGCCATGAAAAATGGCCGCTG	240
QY	241	GAGGAGATCAACAAAGTCGAGATGTGCTGCCGGGCTGCGGCTGGGCTAGCAGCTCTTT	300
Db	241	GAGGAGATCAACAAAGTCGAGATGTGCTGCCGGGCTGCGGCTGGGCTAGCAGCTCTTT	300
QY	301	GATAGCTGCTCGGAGCCTTGTTGGCATGAAGCCAGCCTATATTTCTTGCCCAAGCA	360
Db	301	GATAGCTGCTCGGAGCCTTGTTGGCATGAAGCCAGCCTATATTTCTTGCCCAAGCA	360
QY	361	GGCAGCGGGAATGCGCCCTACTGACACTACAGCACTACAGCACTACAGCACTACAGCACT	420
Db	361	GGCAGCGGGAATGCGCCCTACTGACACTACAGCACTACAGCACTACAGCACTACAGCACT	420
QY	421	GTGATGCGGCGCCCATCTGTCAGAGCTCGCCATGTGTACCGGCAAGTTCTTACGCTTTC	480
Db	421	GTGATGCGGCGCCCATCTGTCAGAGCTCGCCATGTGTACCGGCAAGTTCTTACGCTTTC	480
QY	481	CTCATGCGCCCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	540
Db	481	CTCATGCGCCCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT	540
QY	541	CCCTGCTTCTTCTGCACTGCGCAGCGACCTGTGCAAGCTGACAGCGCCGCGGAGCTT	600

Db	541	CCGTCCTTTCCGACACGCGCCAGAGACCGGTGCACGTACGTACAGGCGCGCGGAGGCTG	600
Oy	601	CTGACAGAGATTCGGGCTGGAACTGGGTGTGGCCGCTTGGGACGACGACGAGATGACGGCCG	660
Db	601	CTGACAGAGATTCGGGCTGGAACTGGGTGGCCGCTTGGGACGACGACGAGATGACGGCCG	660
Oy	661	CAGGGCCGTGAGCATCTTCTGGGCCCTGTGGCCGCGGACGCGGCGCATCTGCATCGCGACAG	720
Db	661	CAGGGCCGTGAGCATCTTCTGGGCCCTGTGGCCGCGGACGCGGCGCATCTGCATCGCGACAG	720
Oy	721	GGCCTGGTGGCCGCTGGCCCGTGGCATGACTGGCGGCTGGGAAAGGTGAGACGCTCTG	780
Db	721	GGCCTGGTGGCCGCTGGCCCGTGGCATGACTGGCGGCTGGGAAAGGTGAGACGCTCTG	780
Oy	781	CACCAAGTGAACACGAGACACGTGCAGAGTGTCTCTGTGGCCCTGGTCAGCCGCC	840
Db	781	CACCAAGTGAACACGAGACACGTGCAGAGTGTCTCTGTGGCCCTGGTCAGCCGCC	840
Oy	841	CACGCGCTCTTCAACTACACATACAGACAGGAGCTCTGCGCCCAAGGTGTGGTGGCAGC	900
Db	841	CACGCGCTCTTCAACTACACATACAGACAGGAGCTCTGCGCCCAAGGTGTGGTGGCAGC	900
Oy	901	GAGGCGTGGGCGACCTCTGACCTGTGCATGTGGGGCTCCCGGGATGGCCAGATGGGCGACG	960
Db	901	GAGGCGTGGGCGACCTCTGACCTGTGCATGTGGGGCTCCCGGGATGGCCAGATGGGCGACG	960
Oy	961	GTGCTTGAGCTTCCTCCAGAGGGGTGGCCACAGCTGCACAGATTCGCCCAAGTGTGAAGAG	1020
Db	961	GTGCTTGAGCTTCCTCCAGAGGGGTGGCCACAGCTGCACAGATTCGCCCAAGTGTGAAGAG	1020
Oy	1021	CACCTGGGCTTGGGACACCGGCTTCTGCTCTGCCCTGGGCGAGAGGAGCAGGCT	1080
Db	1021	CACCTGGGCTTGGGACACCGGCTTCTGCTCTGCCCTGGGCGAGAGGAGCAGGCT	1080
Oy	1081	CTGAGAGAGAGCATGTGTGGGGCCAGACGCTGCCGACATGTGATCAGCTGATCAGCTGACAG	1140
Db	1081	CTGAGAGAGAGCATGTGTGGGGCCAGACGCTGCCGACATGTGATCAGCTGATCAGCTGACAG	1140
Oy	1141	GTGACGCGAGGGCTAAATCACCACACAGAGTCTGTCTGTACGCACTGTGTATAGCGTG	1200
Db	1141	GTGACGCGAGGGCTAAATCACCACACAGAGCTTCTGTCTGTACGCACTGTGTATAGCGTG	1200
Oy	1201	GCCAGAGGCGCTGCACACACACTCTTCAAGTGAAGGCTGAGGGTGGCCCGCGAGAGACCC	1260
Db	1201	GCCAGAGGCGCTGCACACACACTCTTCAAGTGAAGGCTGAGGGTGGCCCGCGAGAGACCC	1260
Oy	1261	GTGAAGCCCTGTGGCAGCTCTGTGGAAGCATGTATCAACTGACCTTCCACTGGGCGGGCTG	1320
Db	1261	GTGAAGCCCTGTGGCAGCTCTGTGGAAGCATGTATCAACTGACCTTCCACTGGGCGGGCTG	1320
Oy	1321	CCGCTGGGGTTCGACACACACCGGAAACGTGTGACATGTGAATTAAGACTGAAAGTGTGGTG	1380
Db	1321	CCGCTGGGGTTCGACACACACCGGAAACGTGTGACATGTGAATTAAGACTGAAAGTGTGGTG	1380
Oy	1381	TGGCAGAGGGCTCAGTGCACAGGCTTCACAGACGTGGGAGTTCACGCGAGCTTCAGAGCA	1440
Db	1381	TGGCAGAGGGCTCAGTGCACAGGCTTCACAGACGTGGGAGTTCACAGCGACGCTTCAGAGCA	1440
Oy	1441	GAGCGCTGAAGATCGGCTGGCACACGTGTGACAAC-----1476	
Db	1441	GAGCGCTGAAGATCGGCTGGCACACGTGTGACAAC-----1476	
Oy	1477	-----CAGAAAGCCGGTGTCGGGTGTGGCGGCAATGGCAGAGAGGCGCAGGTGGC	1527
Db	1501	GCGTGTGCGAGAAAGCCCGGTGTGGCGGCTGTGGCGGCAATGGCAGAGAGGCGCAGGTGGC	1560
Oy	1528	CGGGTCAAGGGGTTCACCTCTCTCTACAGATGTGTGTGAGCTGCAGAGCGGCGAGCTAC	1587
Db	1561	CGGGTCAAGGGGTTCACCTCTCTCTCTACAGATGTGTGTGAGCTGCAGAGCGGCGAGCTAC	1620
Oy	1588	CGGGAAGAACCCAGACGACATCGGCTGACACTTTTGTGGCCAGAGATGATGGTCCCCGAG	1647
Db	1621	CGGGAAGAACCCAGACGACATCGGCTGACACTTTTGTGGCCAGAGATGATGGTCCCCGAG	1680

SOURCE	human.
ORGANISM	Homo sapiens
Eukaryote:	Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia;	Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1
AUTHORS	Patereson, C., Tribouley, C.M., Yao, M.G., Griffin, J.A., Thornton, M., Lu, Y., Kallikch, D.A., Gendhi, A.R. and Au-Young, J.
TITLE	G-protein coupled receptors
JOURNAL	Patent: WO 0190359-A 24 29-NOV-2001;
FEATURES	Incyte Genomics, Inc. (US)
source	Location/Qualifiers 1..2739
BASE COUNT	/organism="Homo sapiens" /db_xref="taxon:9606" Mammalia="Incyte ID No: 7475208CBI"
ORIGIN	415 a 951 c 872 g 501 t
Query Match	81.7%; Score 2090.6; DB 6; Length 2739; Best Local Similarity 91.4%; Pred. No. 4.2e-270; Matches 2455; Conservative 0; Mismatches 19; Indels 213; Gaps 14;
OY	1 ATGCTGGGCCCTGCTGTCTCTGGGACCTCAAGCCCTGTTCCTGCACACCCTGGACGGGG 60
Db	1 ATGCTGGGCCCTGCTGTCTCTGGGACCTCAAGCCCTGTTCCTGCACACCCTGGACGGGG 60
OY	61 GCCCATATGTGCTCTGTCAAGCAACTTGAATGAAGGGGAACTACGTCTGGGGGGCTG 120
Db	61 GCCCATATGTGCTCTGTCAAGCAACTTGAATGAAGGGGAACTACGTCTGGGGGGGGCTG 120
OY	121 TTCCCCCTGGGGGAGCGCGAGGAGGCTGAGCCCTCCGACAGCGGACAGCGCCAGACCCCT 180
Db	121 TTCCCCCTGGGGGAGCGCGAGGAGGCTGAGCCCTCCGACAGCGGACAGCGCCAGACCCCT 180
OY	181 GTGTGCACCAAGTTCTCTCAAAGCGGCTGCTTGGGGCATGTGGCCATGAAAATGACCCGTG 240
Db	181 GTGTGCACCAAGTTCTCTCAAAGCGGCTGCTTGGGGCATGTGGCCATGAAAATGACCCGTG 240
OY	241 GAGGAGATCAACAAGAAGCGGATGTGCTGCCCGGGGCTGGCGCTGGGCTAGAACCTCTTT 300
Db	241 GAGGAGATCAACAAGAAGCGGATGTGCTGCCCGGGGCTGGCGCTGGGCTAGAACCTCTTT 300
OY	301 GATACGTGCTGGAGCCTGTGTGGCCATGAAGCCCAAGCCTCATGTCTCTGGCCAAGGCA 360
Db	301 GATACGTGCTGGAGCCTGTGTGGCCATGAAGCCCAAGCCTCATGTCTCTGGCCAAGGCA 360
OY	361 GGCAACCCCGGACNATGGCCGCTACTGCAATFACAGGCAFTACAGCCCGCGTGTGGCT 420
Db	361 GGCAACCCCGGACNATGGCCGCTACTGCAATFACAGGCAFTACAGCCCGCGTGTGGCT 420
OY	421 GTCATCGGGGCCCACTGCTGTAAGAGTCCGCAATGATCACCGGGAATTTCTTAGCTTTCTC 480
Db	421 GTCATCGGGGCCCACTGCTGTAAGAGTCCGCAATGATCACCGGGAATTTCTTAGCTTTCTC 480
OY	481 CTCATGCCC----- 489
Db	481 CTCATGCCC----- 489
OY	490 ----- -CAGGTACAGTACGGGTCTAGCATMGAG 516
Db	541 GCCCCTGTGTACAGAGATGCTCTTGGGCCCTTGCAAGGTACGTTACGGTGTCTAAGCATGGAG 600
OY	517 CTGCTGAGCGCCCGGGAGACCTTCCCTCTCTTCCGACACCGTCCCAAGCAGACCGTGTG 576
Db	601 CTGCTGAGCGCCCGGGAGACCTTCCCTCTCTTCCGACACCGTCCCAAGCAGACCGTGTG 660
OY	577 CAGGTGAGCGCCCGCGGAGAGTGTCTGCAAGAGTTTCGGTGGCAATGGTGGGCCCGCTTG 636
Db	661 CAGGTGAGCGCCCGCGGAGAGTGTCTGCAAGAGTTTCGGTGGCAATGGTGGGCCCGCTTG 720
OY	637 GGCACGACGACGATAGCGGCGGAGGAGCCGACATCTTTCGGCCCTGGCCGCGGCA 696
Db	721 GGCACGACGACGATAGCGGCGGAGGAGCCGACATCTTTCGGCCCTGGCCGCGGCA 779

[illegible]

OY	1744	CTGTGTCGAGCGCTTTTGGGGGCTGTGCTCAACATCGGGAGACAGCCACTGGTTAGAGCC	1803
Db	1777	CTTGTGTGGCTGTGGGCTGTTCGTACACATCGGGAGACGCCACTGGTTAGAGCC	1836
OY	1804	TGCGGGGGGGCCCTTGAGCCCTGTCTTTTGGCCCTGTGTGATCCCTGGAGCCTGTGCTCAGCGTC	1863
Db	1837	TGCGGGGGGGGGCCCTTGAGCCCTGTCTTTTGGCCCTGTGTGATCCCTGGAGCCTGTGCTCAGCGTC	1896
OY	1864	CTTCTGTTCCTTGCGCAAGCCAGCCCTGCGGATGCTGTGGCCAGACAGCCCTGTTCAC	1923
Db	1897	CTTCTGTTCCTTGCGCAAGCCAGCCCTGCGGATGCTGTGGCCAGACAGCCCTGTTCAC	1956
OY	1924	CTCCCGCTACAGGGGTGGCTGTAGCAGACCTTCCTGACAGGGGGGAGATCTTGTGGGAG	1983
Db	1957	CTCCCGCTACAGGGGTGGCTGTAGCAGACCTTCCTGCTGACAGGGGGGAGATCTTGTGGGAG	2016
OY	1984	TCAGAACTGCGCTGTAGTGGGGCAGACCGGCTGATGTGCTGCTCGGGGGGCGCTTGAGCC	2043
Db	2017	TCAGAACTGCGCTGTAGTGGGGCAGACCGGCTGATGTGCTGCTCGGGGGGCGCTTGAGCC	2076
OY	2044	TGGCTGTGTGTCTCTGTGGCCATGCTGTGTGAGAGTTCGCACTGTGCATCTGGTACTGTGGT	2103
Db	2077	TGGCTGTGTGTCTCTGTGGCCATGCTGTGTGAGAGTTCGCACTGTGCATCTGGTACTGTGGT	2136
OY	2104	GCCCTTCCGCGGAGAGTGGTGAC--GGACTGGACATGCTGCCACGAGAGGCCCTGATGCA	2162
Db	2137	GCCCTTCCGCGGAGAGTGGTGAC--GGACTGGACATGCTGCCACGAGAGGCCCTGATGCA	2196
OY	2163	CTGCGGCACACGCTCTGGGTAGCTTGGCGCTAGCGCAGCGCACCAATGCCACGCTGGC	2222
Db	2197	CTGCGGCACACGCTCTGGGTAGCTTGGCGCGCTAGCGCAGCGCACAC--ATGCGACGCTGGC	2255
OY	2223	CTTTCCTGAGCTTC--CTGGGCACATTCCTGAGTTCGAGGAGCCAGCCGGGGCTGTCTCAACCGTG	2281
Db	2256	CTTTCCTGAGCTTCCTGAGGACTTTCCTGTGTGCGAGCCAGCGGGGCGCTCAACACCGTG	2315
OY	2282	CC---CGTGGCCTTACACTTGGCATGTGAGGCTACTTTCATTCAC--CTGGGCTTC--CTTTGT	2336
Db	2316	CCAGCTGGGCTTACACATTTTGGCAATGCTGTGGCTACTTTCATTCACACTGGGCTCTCACTTGT	2375
OY	2337	GCCCTCTCTGGC--CAATGTGACAGTGTGT--CCTCAGGCCCGCGCTGTGACAGATGGGCGCCT	2393
Db	2376	GCCCTCTCTGGCACATGTGACAGTGTGTGTACTCAAGGCCCGACGCTGTGACAGATGGGCGCCT	2435
OY	2394	CTGTCTGTGTCTCAGGGGATCTGTGGTGCCTTCCACCGGCCAGGTGTACTCTGTCTAT	2453
Db	2436	CTGTCTGTGTCTCAGGGGATCTGTGGTGCCTTCCACCGGCCAGGTGTACTCTGTCTAT	2495
OY	2454	GGGCGACCGAGGGCTCAACACCCCGAGTTCCTTCTGGGAGGGGGCCCTGGGGATGCC--C	2512
Db	2496	GGGCGACCGAGGGCTCAACACCCCGAGTTCCTTCTGGGAGGGGGCCCTGGGGATGCCAC	2555
OY	2513	AAGGCGAAGTATGAGGGGAACAGGAATCTACAGGGGAACATATAGTGA	2559
Db	2556	AAGGCGAAGTATGAGGGGAACAGGAATCTACAGGGGAACATATAGTGA	2602

	RESULT 7	
AB065647	LOCUS	
	DEFINITION	AB065647 4256 bp DNA linear PRI 23-JUL-2002 Homo sapiens gene for seven transmembrane helix receptor, complete cds, isolate:CBRC7TM_210.
ACCESSION		AB065647
VERSION		AB065647.1 GI:21928568
KEYWORDS		.
SOURCE		Homo sapiens (isolate:CBRC7TM_210) DNA.
ORGANISM		Homo sapiens
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
		Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE		1 Suwa,M., Sato,T., Okouchi,I., Arita,M., Futami,K., Matsumoto,S., Tsutsumi,S., Aburatani,H., Asai,R. and Akiyama,Y. Genome-wide discovery and analysis Of human seven transmembrane
AUTHORS		
TITLE		

JOURNAL helix receptor genes
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 4256)
TITLE Suwa,M.
JOURNAL Direct Submission

Submitted (11-JUL-2001) Makiko Suwa, Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST); 2-41-6 Aomi Koto-ku, Tokyo 135-0064, Japan (E-mail:m-suwa@aist.go.jp, URL:http://www.cbrc.jp/, Tel:81-3-3599-8080, Fax:81-3-3599-8081)

This sequence is a seven transmembrane helix receptor candidate predicted from the whole human genome sequences using our automated system that contains programs of gene finding(GeneDecoder), sequence search, motif-domain assignment and transmembrane helix prediction.

And the sequence is submitted by the collaborative project between [Computational Biology Research Center (CBRC), National Institute of Advanced Industrial Science and Technology (AIST)] and [Genome Science Division, Research Center for Advanced Science and Technology (KCAST), University of Tokyo].

Location/Qualifiers
1..4256

FEATURES
source
/organism="Homo sapiens"
/isolate="CBRC/TM_210"
/db_xref="taxon:9606"
/chromosome="1"
join(201..391,493..793,879..1661,1776..1979,2114..223361..3315,3713..3890,3979..4056)
/codon_start=1
/evidence="not_experimental"
/product="seven transmembrane helix receptor"
protein_id="BAC05873.1"
/db_xref="GI:21928569"
/translation="MLGPSPAYGLSLMALHPGTGAPICLSQOLRMKGDIYVIGLTFPLGSAAEPAGSRTPSSPVCTRRSDNSNGILMALMKAAVEINNKSLPLRGXDLPFTCEAEVVMKRPSEIMFLAKRAGSDIAKYNTGYOORVLAIVGPSSELAMTKGRFEE LMPDYSGASMEILLARETPEFSFRITRYPSDRVQLTAAILLEPGRNMYVAALSDDTY GRGDLSTFSAALARGICITALEGLVLPTRADDSVLGVVDLVHOVNOSVOVALLFNSS VHAHALFNYSISSRLSPKPWVAEAMLTDSLVLGFGMAOMGVIAELDROGAHLFEFS POYKTHTLATIDPAFCSSALGEREGLEDIYVGRCPCDCITLOANSACLINHQTFS YVAAYVSVALATDHNTLOCNASGCCAOPDPFWKMOLLNNMTLFHWGLPLRFDSGNV DMEVDIKLMWGCSVPLRHVDGRNGSIIRERLKIRMTSHONOKPYRSKOOCOEQOV RRVKGFHSCTCYDCERAGSYRQRPDICTACTGCODMSPPERSRCTRRRSRLAMEB PAVYLILSLISLAGLVLAALGLTVHRHDSLVAGSPPLACFGLVCLGLSVLFL PGQSPAPACLAQPLSHLPITGCLSTETLDALETVESELPLSWADRSLCGLRGMWM LVNLIAMLEVVALCTWLVAFPEPVYDMMLPTREALVHCRTBSWSFGLAHNATNL AFCLFGLFTELRSPGPCYNRARGLTFAMLNVFIITWVSFVLNAVYVLPAVOMGALL LVCVGLIAALFPRCLYILMRQPLINRDEPIFGGPDGAGONGNTGNNGKHAPLLI HPAAHHSSGRGGCGTOAYAGCRARRGCGQLPLPFAOWYDSIMYTSTRDRADQILT PARQAHTGSPVRYPGSAFPSRA"

BASE COUNT 707 a 1445 c 1370 g 734 t
ORIGIN

Query Match 76.2%; Score 19.9; DB 9; Length 4256;
Best Local Similarity 82.0%; Pred. No.3e-251;
Matches 2559; Conservative 0; Mismatches 0; Indels 560; Gaps 5;

DY 1 ATGCTGGGGCCCTGCTCTGCTGGGCTCAAGCCCTGTCCTCTCACCTTGGAACGGGG 60
DB 201 ATGCTGGGGCCCTGCTCTGCTGGGCTCAAGCCCTGTCCTCTCACCTTGGAACGGGG 260

DY 61 GCCCATTTGTGCTGTCAAGCAACAATTAGATGAAGGGGGACTACCTCTGGGGGGCTG 120
DB 261 GCCCATTTGTGCTGTCAAGCAACAATTAGATGAAGGGGGACTACCTCTGGGGGGCTG 320

DY 121 TTCCCCCTGGGCGAGGCGGAGAGGCTGACCCTCGCGACCGGACACGGCCCAAGACCCCT 180
DB 321 TTCCCCCTGGGCGAGGCGGAGAGGCTGACCCTCGCGACCGGACACGGCCCAAGACCCCT 380

DY 181 GTGTGCA----- 187
DB 381 GTGTGCAAGGTACAGAGGTGGGACGGCGTGGGGTCCAGGGTACCAAGGTCTGGG 440

QY 188 -----CAGGTTCTCT 199
DB 441 GTGCTCTGAGCTGGGGCCGAGGTGGCCATCTGCGGTTCTGTGTGCCCCCAGGTTCTCT 500
QY 200 CAAAGGCGCTGCTCTGGGCACTGGCCATGAAATGCGCGTGGAGAGATCAACAACAGT 259
DB 501 CAAAGGCGCTGCTCTGGGCACTGGCCATGAAATGCGCGTGGAGAGATCAACAACAGT 560
QY 260 CGGATCTGCTGGCGGGCTGGCGCGTGGGCTAGGACTCTTTGATACGTCGTGGAGCTG 319
DB 561 CGGATCTGCTGGCGGGCTGGCGCGTGGGCTAGGACTCTTTGATACGTCGTGGAGCTG 620
QY 320 TGGTGGCCATGAAAGCCGACCTCATGTTCTTGGCCAAAGGAGGAGCCGCAATCGCG 379
DB 621 TGGTGGCCATGAAAGCCGACCTCATGTTCTTGGCCAAAGGAGGAGCCGCAATCGCG 680
QY 380 CCTACTGCACTACACGAGATACAGCCCGTGTGCTGGCTGTCAATCGGGCCCACTG 439
DB 681 CCTACTGCACTACACGAGATACAGCCCGTGTGCTGGCTGTCAATCGGGCCCACTG 740
QY 440 CAGAGCTGGCCATGCTACCGCGGAGTTCTTCAAGCTTCTCTCATGCGC----- 489
DB 741 CAGAGCTGGCCATGCTACCGCGGAGTTCTTCAAGCTTCTCTCATGCGCAGTGGCG 800
QY 490 ----- 489
DB 801 CCCCCACATCACCCACCCACCCCTGCCCCGTTGGAGGCCCTGTGTGACAGAGAT 860
QY 490 -----CAGTCACTAGCGTGTGATGAGAGTCTGAGCGCCGGAG 534
DB 861 GCGTCTGGCCCTTTCAGAGTCACTAGTACGATGAGAGTCTGAGCGCCGGAG 920
QY 535 ACCTTCCCTCTCTTCTTCCGACGCTGGCCAGCGACCGTGTGACATGAGCGCCCGCG 594
DB 921 ACCTTCCCTCTCTTCTTCCGACGCTGGCCAGCGACCGTGTGACATGAGCGCCCGCG 980
QY 595 GAGCTGTGACAGAGTTCGGCTGGAATGCGTGGCGCCCTTGGAGGAGCAGCAGAGTAC 654
DB 981 GAGCTGTGACAGAGTTCGGCTGGAATGCGTGGCGCCCTTGGAGGAGCAGCAGAGTAC 1040
QY 655 GGGCGGAGGCGCTGAGATCTTCTCGGCGCTGGCGCGGAGGAGGAGATCTGATCGCG 714
DB 1041 GGGCGGAGGCGCTGAGATCTTCTCGGCGCTGGCGCGGAGGAGGAGATCTGATCGCG 1100
QY 715 CACGAGGCGCTGTGCTCCGCTGCGCCGCTGCGATGACTGCGGCGTGGGAGGATGAGAG 774
DB 1101 CACGAGGCGCTGTGCTCCGCTGCGCCGCTGCGATGACTGCGGCGTGGGAGGATGAGAG 1160
QY 775 GTCTGTGACACAGGTGACACAGAGCGTGTGAGTGTGCTGTGCTGCGCTCGTGAC 834
DB 1161 GTCTGTGACACAGGTGACACAGAGCGTGTGAGTGTGCTGTGCTGCGCTCGTGAC 1220
QY 835 GCGGCGGACGCGCTTCTTCACTACAGCAGTACAGCAGAGCTTCCGCCAAGTGTGGGTG 894
DB 1221 GCGGCGGACGCGCTTCTTCACTACAGCAGTACAGCAGAGCTTCCGCCAAGTGTGGGTG 1280
QY 895 GCGACGAGGCGCTGTGCTTACCTGTGACTGTGATGGGCTGCCGGATGGCCAGATG 954
DB 1281 GCGACGAGGCGCTGTGCTTACCTGTGACTGTGATGGGCTGCCGGATGGCCAGATG 1340
QY 955 GGCAGGTGTGCTGCTTCTCCAGAGGGGTGCCAGGTGACAGAGTTCGCCAGTACGTG 1014
DB 1341 GGCAGGTGTGCTGCTTCTCCAGAGGGGTGCCAGGTGACAGAGTTCGCCAGTACGTG 1400
QY 1015 AAGAGGCACTGGCGCGCTGGCCACGACCGGCTTCTGCTGTGCGCTGGGAGAGGAG 1074
DB 1401 AAGAGGCACTGGCGCGCTGGCCACGACCGGCTTCTGCTGTGCGCTGGGAGAGGAG 1460
QY 1075 CAGGCTCTGAGAGAGAGTGTGGTGGGAGGCTGCCGAGTGTGATGATGATCAACGCTG 1134
DB 1461 CAGGCTCTGAGAGAGAGTGTGGTGGGAGGCTGCCGAGTGTGATGATGATCAACGCTG 1520
QY 1135 CAGAACTGTAGAGGCAAGGCTAAATCACACAGAGAGTTCGTGTGTACGAGCTGTGTAT 1194

DB 1521 CAGAACTGTAGAGGCAAGGCTAAATCACACAGAGAGTTCGTGTGTACGAGCTGTGTAT 1580
QY 1195 AGCGTGGCCAGGCGCTGACACACTTTCATGATGATGATGATGATGATGATGATGATGAT 1254
DB 1581 AGCGTGGCCAGGCGCTGACACACTTTCATGATGATGATGATGATGATGATGATGATGAT 1640
QY 1255 GACCCCGTGAAGGCGCTG----- 1271
DB 1641 GACCCCGTGAAGGCGCTGAGGATGAGCCCGGAGATGGGGGTGTGTCTCTGTGATGT 1700
QY 1272 ----- 1271
DB 1701 GCGCCAGGCAACGAGGCGACGCGCTGAGCTGAGAGTGGTGGGCTGAGCCCG 1760
QY 1272 -----GACGCTCTGGAGACATGTACAACTGACCTTCCAGCTGGCGGCTG 1320
DB 1761 GTCCCGCGCGCAGCTCTGAGAACTGTACAACTGACCTTCCAGCTGGCGGCTG 1820
QY 1321 CCGCTGGGCTGACAGAGGCGGAAAGTGGACATGAGATGAGCTGAGAGCTGTGGGTG 1380
DB 1821 CCGCTGGGCTGACAGAGGCGGAAAGTGGACATGAGATGAGCTGAGAGCTGTGGGTG 1880
QY 1381 TGGCAGGCGTGAAGTCCAGGCTCCAGCAGCTGGGCAAGTTCACAGCAGCTTACAGACA 1440
DB 1881 TGGCAGGCGTGAAGTCCAGGCTCCAGCAGCTGGGCAAGTTCACAGCAGCTTACAGACA 1940
QY 1441 GAGCGCTGAGAGTCCGCTGGCAGACGCTGTGACAC----- 1476
DB 1941 GAGCGCTGAGAGTCCGCTGGCAGACGCTGTGACAC----- 2000
QY 1477 ----- 1476
DB 2001 CAGGCTGCGCGTGTGTAGCCCCCGCGGAGGCGAGGCTGGGGGTGGGGCGCTTCCAG 2060
QY 1477 -----CAGAGCCCG 1486
DB 2061 TCTCCGTGGGATGCCAGCGAGCAGAGCAGACCCAGGCTGTGGCGAGAGCCCG 2120
QY 1487 TGTCCCGGCTGCGGCGGAGTGTCCAGAGAGGAGTGTGCGCGCGGCTGCAAGGGCTTCACT 1546
DB 2121 TGTCCCGGCTGCGGCGGAGTGTCCAGAGAGGAGTGTGCGCGCGGCTGCAAGGGCTTCACT 2180
QY 1547 CCTGCTGTACAGTGTGTGACTGCGAGGCGGAGGAGTGTGCGCGGAGTGTGCGCGGAGTGTG 1597
DB 2181 CCTGCTGTACAGTGTGTGACTGCGAGGCGGAGGAGTGTGCGCGGAGTGTGCGCGGAGTGTG 2240
QY 1598 ----- 1597
DB 2241 GCGCTTCCCGGAGGCGGCGGAGTGTGAGAGCAGCAGGAGGAGTGTGCGCAAGTCTGACT 2300
QY 1598 -----CAG 1600
DB 2301 CTGAGACAGAGCCACAGAGGTGACAGAGAACACACACGCGCTTCTCTCTACAG 2360
QY 1601 ACAGATGCGCTGTGACCTTTTGTGCGAGATGATGTGTCCCGGAGGAGACAGCAGCT 1660
DB 2361 ACAGATGCGCTGTGACCTTTTGTGCGAGATGATGTGTCCCGGAGGAGACAGCAGCT 2420
QY 1661 GCTTCCGCGAGTGTGCTGCTGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1720
DB 2421 GCTTCCGCGAGTGTGCTGCTGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2480
QY 1721 TGTGCTGTGAGGCTGTGCGGCTGTGCGCTGTGCGCTGTGCGCTGTGCGCTGTGCGCTGTG 1780
DB 2481 TGTGCTGTGAGGCTGTGCGGCTGTGCGCTGTGCGCTGTGCGCTGTGCGCTGTGCGCTGTG 2540
QY 1781 GGGACAGCCCACTGTGTAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1840
DB 2541 GGGACAGCCCACTGTGTAGGCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2600
QY 1841 TGGGCTGTGTGCTGTGAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1900


```

Db 2601 TGGGCGTGTGCTGACAGGCTCCTGTTCCCTGAGCAGCCAGCCCTGCCCCGATGCC 2660
QY 1901 TGGCCAGCAGACCCCTTGTCCACCTCCCGCTCAGCGGCTGCTGAGCACAATCTTCTGCG 1960
Db 2661 TGGCCAGCAGACCCCTTGTCCACCTCCCGCTCAGCGGCTGCTGAGCACAATCTTCTGCG 2720
QY 1961 AGCGGCGCAGAGATCTTGTGAGAGTACAGAACTGCTCTGAGCTGAGGAGAGCCGCTGAGTG 2020
Db 2721 AGCGGCGCAGAGATCTTGTGAGAGTACAGAACTGCTCTGAGCTGAGGAGAGCCGCTGAGTG 2780
QY 2021 GCTTCCCTGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2080
Db 2781 GCTTCCCTGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 2840
QY 2081 CACTGTGACCTGTGTACCTGTGTGCTTCCCGCGAGAGTGTGACGAGTGTGACATGTC 2140
Db 2841 CACTGTGACCTGTGTACCTGTGTGCTTCCCGCGAGAGTGTGACGAGTGTGACATGTC 2900
QY 2141 TGGCCAGCAGAGGCGCTGTGTGCTGCTGCGCAGACAGCTCTTGGGTCAGCTTTCGGCTAGCGC 2200
Db 2901 TGGCCAGCAGAGGCGCTGTGTGCTGCTGCGCAGACAGCTCTTGGGTCAGCTTTCGGCTAGCGC 2960
QY 2201 AGCGCACCAATGACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2260
Db 2961 AGCGCACCAATGACAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3020
QY 2261 AGCGGCGCTGTACACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2320
Db 3021 AGCGGCGCTGTACACACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3080
QY 2321 CTTGGGCTCTCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2380
Db 3081 CTTGGGCTCTCTTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3140
QY 2381 AGATGGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2440
Db 3141 AGATGGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3200
QY 2441 GTTACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2500
Db 3201 GTTACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3260
QY 2501 CTGGGAGTGGCCAGAGGCGAGAAATGAGGAGAAACAGAGAAATGAGGAGAAATGAGGAG 2559
Db 3261 CTGGGAGTGGCCAGAGGCGAGAAATGAGGAGAAACAGAGAAATGAGGAGAAATGAGGAG 3319

```

```

RESULT 8
AL139287 84445 bp DNA linear PRI 26-JUN-2002
LOCUS Human DNA sequence from clone RP5-89003 on chromosome 1, complete
DEFINITION sequence.
ACCESSION AL139287
VERSION AL139287
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
1 (bases 1 to 84445)
REFERENCE 1
AUTHORS Leonamornlert,D.
TITLE Direct Submission
JOURNAL Submitted (26-JUN-2002) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquerry@sanger.ac.uk; Clone requests: clonerequests@sanger.ac.uk;
On Jun 27, 2002 this sequence version replaced gi:21615561.
COMMENT ----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquerry@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.

```

Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality > 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/C.elegans/wormpep>

This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chrl>

RP5-89003 is from the library RPCI-5 constructed by the group of Pieter de Jong. For further details see <http://www.choi1.org/bacpac/home.htm>

VECTOR: pCIPAC2.

FEATURES

source

```

1..84445
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP5-89003"
/clone_1lb="RPCI-5"
BASE COUNT 17111 a 26105 c 26413 g 14816 t
ORIGIN

```

```

Query Match 76.2%; Score 1949; DB 9; Length 84445;
Best Local Similarity 82.0%; Pred. No. 1.6e-251;
Matches 2559; Conservative 0; Mismatches 0; Indels 560; Gaps 5;

```

```

QY 1 ATGCTGGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 60
Db 31299 ATGCTGGGCGCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 31358
QY 61 GCCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120
Db 31359 GCCCATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 31418
QY 121 TTCCCTGCGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 180
Db 31419 TTCCCTGCGGCGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCT 31478
QY 181 GTGTGCA----- 187
Db 31479 GTGTGCAAGGATACAGAGGTGGAGCGGCTGGGTGGGGTCAAGGTGACAGGTGTGGG 31538
QY 188 -----CCAGGTTTCCT 199
Db 31539 GTGTCTGTGAGCTGGGCGGAGGTGGCCATCTGCGGTTCTGTGTGGGCCAGAGTTTCCT 31598
QY 200 CAAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 259
Db 31599 CAAGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 31658
QY 260 CGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 319
Db 31659 CGGATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 31718
QY 320 TGTGGGCAATGAAGCCAGGCTCATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 379
Db 31719 TGTGGGCAATGAAGCCAGGCTCATGTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 31778
QY 380 CTTACTGCACTACGACGATACAGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 439

```


D	31779	CCACTGCACATACACGCAAGTACACAGCCCCGCTGCTGCTCATCGGGCCCCACTGCT	31838
OY	440	CAGAGCTGCACATGATACACGGCAAGTCTTCAGCTTCTTCATGCTCC-----	489
D	31839	CAGAGCTGCACATGATACACGGCAAGTCTTCAGCTTCTTCATGCTCC-----	31898
OY	490	-----	489
D	31899	CCCCCACCACATACCCACCCACCAGCCCTGCCCCGTGGAGCCCCCTGTCAAGAGAT	31958
OY	490	-----CAGGTCAAGTACAGGTGTCTAGCATGAGACTCTGTGCCCCGGAG	534
D	31959	GCCCTTGAGCCCTTGACAGGTGACTAGAGGTGTACATGAGAGCTCTGTAGCCCCGGAG	32018
OY	535	ACCTTCCTCCCTCTTCTTCGCGACACGCTGCCACAGCACTGTGTACGTACAGCCCGCG	594
D	32019	ACCTTCCTCCCTCTTCTTCGCGACACGCTGCCACAGCACTGTGTACGTACAGCCCGCG	32078
OY	595	GAGCTGTGCAGAGATTTGGGTGTGAAGTGGGGGGCCCTGGGGAGGACAGCAGATAC	654
D	32079	GAGCTGTGCAGAGATTTGGGTGTGAAGTGGGGGGCCCTGGGGAGGACAGCAGATAC	32138
OY	655	GGCCGGCAGGGCTTGAGCATCTTCTGGCCCTGGCCGGCACGCGGCATCTGCATCGCG	714
D	32139	GGCCGGCAGGGCTTGAGCATCTTCTGGCCCTGGCCGGCACGCGGCATCTGCATCGCG	32198
OY	715	CACGAGGGCCGTGGTGGCCCTGGCCGTGCCATGTACTGGGGCTGGGGAAGGTGCAGAG	774
D	32199	CACGAGGGCCGTGGTGGCCCTGGCCGTGCCATGTACTGGGGCTGGGGAAGGTGCAGAG	32258
OY	775	GTCTGTGCACAGATGAACAGACAGACGCTGTGAGGTGTGTCTTGTGCTCTCCGTGCAC	834
D	32259	GTCTGTGCACAGATGAACAGACAGACGCTGTGAGGTGTGTCTTGTGCTCTCCGTGCAC	32318
OY	835	GCCGCCACAGGCTTTTCACTACAGCATCAGACAGGCTCTGCCCCAGGTGTGGGTG	894
D	32319	GCCGCCACAGGCTTTTCACTACAGCATCAGACAGGCTCTGCCCCAGGTGTGGGTG	32378
OY	895	GCCACGAGAGGCTGTGCTACCTCTACCTGTGTATGAGGGCTCCCGGCACTGGCCAGATG	954
D	32379	GCCACGAGAGGCTGTGCTACCTCTACCTGTGTATGAGGGCTCCCGGCACTGGCCAGATG	32438
OY	955	GGCAGCGTGTGTGGCTTCTCCAGAGGGGTGCCACAGCTGCACAGATTCCTCCAGTACGTG	1014
D	32439	GGCAGCGTGTGTGGCTTCTCCAGAGGGGTGCCACAGCTGCACAGATTCCTCCAGTACGTG	32498
OY	1015	AAGACGCACTGGCCCTGGCCACGCAACCCGGCTTCTGTGCTCCCTGGCCAGAGGGAG	1074
D	32499	AAGACGCACTGGCCCTGGCCACGCAACCCGGCTTCTGTGCTCCCTGGCCAGAGGGAG	32558
OY	1075	CAGGGTGTGAGAGAGAGGTGTGGGCCAGGCTCCCGCACTGTGTACTGTACACGCTG	1134
D	32559	CAGGGTGTGAGAGAGAGGTGTGGGCCAGGCTCCCGCACTGTGTACTGTACACGCTG	32618
OY	1135	CAGAACGTGAGGGGCAAGGCTTAATATACCAACAGAGCTTCTGTCTACGACTGTGTAT	1194
D	32619	CAGAACGTGAGGGGCAAGGCTTAATATACCAACAGAGCTTCTGTCTACGACTGTGTAT	32678
OY	1195	AGCTGTGGCCACAGGCCCTGTGCAACACTCTTTCAGTGCACAGCCTTCAGAGGCTGCCCGCAG	1254
D	32679	AGCTGTGGCCACAGGCCCTGTGCAACACTCTTTCAGTGCACAGCCTTCAGAGGCTGCCCGCAG	32738
OY	1255	GACCCCGTAGAGCCCTG-----	1271
D	32739	GACCCCGTAGAGCCCTGTGGCAGGTGAGCCCGGAGATGGGGGTGTGCTGTCTCTGCATGT	32798
OY	1272	-----	1271
D	32799	GCCACAGGCCACAGGCAAGGCCACCAACGCTCTAGCTGGAGGTGGCTGGCGGCTCAAGCCC	32855
OY	1272	-----GACGCTCTGTGAGAACATGTATCAACCTGTGACTTCCACGTGGGGGGGCTG	1320
D	32859	GTCCCCCGCCGACACTCTCTGTGAGAAATGTATCAACTGTACCTTCACGTGGGGGGCTG	32918

QY	1321	CCGCTGCAGTTTGACAGCAGCGGAAACGTGGACATGAGACTACGACCTGAAGCTGTGGTGC	1380
Dd	32919	CCGCTGCAGTTTGACAGCAGCGGAAACGTGGACATGAGACTACGACCTGAAGCTGTGGTGC	32978
QY	1381	TGGCAGGGGCTCAAGTCGCCAGGCTCCACAGCAGTGGGGCAGGTTAACGGCGAGCTCCAGGACA	1440
Dd	32979	TGGCAGGGGCTCAAGTCGCCAGGCTCCACAGCAGTGGGGCAGGTTAACGGCGAGCTCCAGGACA	33038
QY	1441	GAGCGCCTGAAGATCCGCTGGGCACACGTCGTGCAAC-----	1476
Dd	33039	GAGCGCCTGAAGATCCGCTGGGCACACGTCGTGCAAC-----	33098
QY	1477	-----	1476
Dd	33099	CAGGCGTCCCCTGTGAGTACC-----CCGCGCACAGGCGCAGCCTGGGGGTGGGGCCGTTTCAG	33158
QY	1477	-----CAGAAGCCCG	1486
Dd	33159	TCTCCCGTGGGAGTAGCCACGCGAGAGAGCAGAACCCAGGCCCTGTGGCGAGAAAGCCCG	33218
QY	1487	TGTCCCCGTGCTTCGGGCGCAGTGCACAGAGAGGCCAAGTGCGCCGGGTCAAGGGGTTTCCACT	1546
Dd	33219	TGTCCCCGTGCTTCGGGCGCAGTGCACAGAGAGGCCAAGTGTGCCGGGTCAAGGGGTTTCCACT	33278
QY	1547	CCTGCGCTACACATGTGTGGACTGAGAGCGGCGGCGACTACCGCGCAAAAC-----	1597
Dd	33279	CCTGCGCTACACATGTGTGGACTGAGAGCGGCGGCGACTACCGCGCAAAAC-----	33338
QY	1598	-----	1597
Dd	33339	CGCCTTCCCGGAGCGGGGTGGGAGCAGCAGAGGGAGGGTCTCTGCCMAATCCTGACT	33398
QY	1598	-----CAG	1600
Dd	33399	CTGAGACAGAGCCCCACAGGGTACAAAGACCAACCCAGCGCCCTTCTCTTCACAG	33458
QY	1601	ACGACATCGCTGCACACCTTTTGTGGCCAGATGAGTGTCCCGGAGCGAAGACACAGCT	1660
Dd	33459	ACGACATCGCTGCACACCTTTTGTGGCCAGATGAGTGTCCCGGAGCGAAGACACAGCT	33518
QY	1661	GCTTCGCGCGCAGTCTCGGTTCCGTGGCATGGGGGAGCGCGGCTGTGCTGCTGCTCTCC	1720
Dd	33519	GCTTCGCGCGCAGTCTCGGTTCCGTGGCATGGGGGAGCGCGGCTGTGCTGCTGCTCTCC	33578
QY	1721	TGCTGCTGAGCCTGGCGCGCTGGGCTGTGTCAGTGTGCTTTGGGGGCTGTGCTTACACATC	1780
Dd	33579	TGCTGCTGAGCCTGGCGCGCTGGGCTGTGTCAGTGTGCTTTGGGGGCTGTGCTTACACATC	33638
QY	1781	GGGACAGCCCATGTGTTCAAGSCCTCGGGGGGCGCCCTGGCCTGTGTTGGCCTGTGTGCC	1840
Dd	33639	GGGACAGCCCATGTGTTCAAGSCCTCGGGGGGCGCCCTGGCCTGTGTTGGCCTGTGTGCC	33698
QY	1841	TGGGGCTGGTTCGCMCACAGTCTGCTTCCCTGGGCCACAGCCCTGGCCGATGGC	1900
Dd	33699	TGGGGCTGGTTCGCMCACAGTCTGCTTCCCTGGGCCACAGCCCTGGCCGATGGC	33758
QY	1901	TGGCCACAGAGCCCTGTTGCCACACTCCGCTCACGGGCTGCTTGAGACACACTTCTCTGC	1960
Dd	33759	TGGCCACAGAGCCCTGTTGCCACACTCCGCTCACGGGCTGCTTGAGACACACTTCTCTGC	33818
QY	1961	AAGCGGCGAGATCTTCGTGGATCAGAACTGCTGTAGCTGGGAGACCGCGCTGAGTG	2020
Dd	33819	AAGCGGCGAGATCTTCGTGGATCAGAACTGCTGTAGCTGGGAGACCGCGCTGAGTG	33878
QY	2021	GCTGCTGGGGGGCCCTGGGCTGGCTGGTGGTGTGCTGTGGCCATGCTGTGTGAGAGTGC	2080
Dd	33879	GCTGCTGGGGGGCCCTGGGCTGGCTGGTGGTGTGCTGTGGCCATGCTGTGTGAGAGTGC	33938
QY	2081	CACGTGACACCTGTGTAACCTGTGTGGCTTCCCGCCGGAAGTGTGTAGAGACTGGCACATGC	2140
Dd	33939	CACGTGACACCTGTGTAACCTGTGTGGCTTCCCGCCGGAAGTGTGTAGAGACTGGCACATGC	33998

[illegible]

```

|||||
Db 72308 CAGAGTCCCAATGGTACACGGGAAGTTCTTCAGCTTCTTCTCATGCCCCAGAGTGGCC 72367
QY 490 ----- 489
Db 72368 CCCCCACATCAACCCACCCACCCAGCCCTGCGGAGCCCTGTGTAGAGATG 72427
QY 490 -----CAGGTACGTACGGTGTAGATGAGTGTGAGCTGTGAGCGCCGGGAGA 535
Db 72428 CCTCTTGGCCCTTGCAGGTACGTACGTGTGATGAGTGTGAGCTGTGAGCGCCGGGAGA 72487
QY 536 CCTTCCCTCTCTCTCTCTCCACCGTGTCCAGGACCGTGTGAGCTGTGAGCGCCGGCGG 595
Db 72488 CCTTCCCTCTCTCTCTCTCCACCGTGTCCAGGACCGTGTGAGCTGTGAGCGCCGGCGG 72547
QY 596 AGCTGTGAGAGATTGGCTGTGAGTGGTGGCGCCCTGGGGACGACGAGAGTACG 655
Db 72548 AGCTGTGAGAGATTGGCTGTGAGTGGTGGCGCCCTGGGGACGACGAGAGTACG 72607
QY 656 GCGGCAAGGCGCTGAGCATCTTCTGCGCCCTGGCGCGGACAGCGGSCATCTGCATGGCG 715
Db 72608 GCGGCAAGGCGCTGAGCATCTTCTGCGCCCTGGCGCGGACAGCGGSCATCTGCATGGCG 72666
QY 716 ACAGAGGCGCTGTGCGCGCTGCGCCGATGACTGCGGCGTGGGGAAGTGCAGAGAG 775
Db 72667 ACAGAGGCGCTGTGCGCGCTGCGCCGATGACTGCGGCGTGGGGAAGTGCAGAGAG 72726
QY 776 TCGTGCACCAAGTGAACCAAGACAGAGTGTGAGTGTGTCTGTCTGCTCCCTGTCACG 835
Db 72727 TCGTGCACCAAGTGAACCAAGACAGAGTGTGAGTGTGTCTGTCTGCTCCCTGTCACG 72786
QY 836 CCGCCACAGCCCTCTTCAACTACAGCATAGCAGCAGGCTCTGCGCCAAAGTGTGGGTG 895
Db 72787 CCGCCACAGCCCTCTTCAACTACAGCATAGCAGCAGGCTCTGCGCCAAAGTGTGGGTG 72846
QY 896 CCAGGAGGCGCTGTGCTGACCTGTGACCTGTGTATGAGGGCTGCCCGGACATGSCCAGATG 955
Db 72847 CCAGGAGGCGCTGTGCTGACCTGTGTATGAGGGCTGCCCGGACATGSCCAGATG 72906
QY 956 GCAGGCTGTGGCTTCCCTCCAGAGGGGTGCCAGGTGTGAGAGTTCGCCAGTACGTGA 1015
Db 72907 GCAGGCTGTGGCTTCCCTCCAGAGGGGTGCCAGGTGTGAGAGTTCGCCAGTACGTGA 72966
QY 1016 AGACGCACTGTGCTGTGCGCACGACCGGCTTGTGTCTGTCCCTGCGGCGAGAGGAGC 1075
Db 72967 AGACGCACTGTGCTGTGCGCACGACCGGCTTGTGTCTGTCCCTGCGGCGAGAGGAGC 73026
QY 1076 AGGCTGTGAGAGAGAGTGTGTGGGCGACGCTGCCCGAGTGTGACTGCATACGCTGC 1135
Db 73027 AGGCTGTGAGAGAGAGTGTGTGGGCGACGCTGCCCGAGTGTGACTGCATACGCTGC 73086
QY 1136 AGAAGTGAAGCGCAGAGGCTAAATCAACACAGAGTTCCTGTCTAAGAGAGTGTGTATA 1195
Db 73087 AGAAGTGAAGCGCAGAGGCTAAATCAACACAGAGTTCCTGTCTAAGAGAGTGTGTATA 73146
QY 1196 GCGTGGCCCGGCGCTGTGACCAACACTCTTCACTGTGCAAGCGCTCAAGGCTCCCGCGCAG 1255
Db 73147 GCGTGGCCCGGCGCTGTGACCAACACTCTTCACTGTGCAAGCGCTCCCGCGCAG 73205
QY 1256 ACCCGGTGAAGCCCTG----- 1271
Db 73206 ACCCGGTGAAGCCCTGAGGTGAGCCCGGAGATGGGGGTGTCTGTCTGTGCAATG 73265
QY 1272 ----- 1271
Db 73266 CCGAGGCAACAGGAGCGGCAACAGCGCTGAGCTGTGAGTGTGGCTGTGACGCGCG 73325
QY 1272 -----GAGCTCTGTGAAGAACATGTACAACTGACCTTCAAGTGTGGCGGCTGC 1321
Db 73326 TCCCGCGCGCGAGCTCTCGGGAACATGTACAACTGACCTTCAAGTGTGGCGGCTGC 73385
QY 1322 GCGTGGGTGAGAGAGCGGAACGTGTGAGATGAGTGTGAGTGTGAGTGTGAGTGTG 1381
|||||

```

```

Db 73386 CGCTGGGTTTCAGACAGCAGCGGAAACGTGTGATGTAGCATCTGAAGTGTGGGTGT 73445
QY 1382 GCGAGGCGTCAAGTGTGCGCAGGCTTCACGACGTGGGCAAGTTCAAGCGACCTTGAGCAG 1441
Db 73446 GCGAGGCGTCAAGTGTGCGCAGGCTTCACGACGTGGGCAAGTTCAAGCGACCTTGAGCAG 73505
QY 1442 AGCGGCTGAAGATCCGCTGTGCAACAGCTGTGCAAC----- 1476
Db 73506 AGCGGCTGAAGATCCGCTGTGCAACAGCTGTGCAACAGCTGTGAGTGTGAGTGTGCC 73565
QY 1477 ----- 1476
Db 73566 AGCGGCGCGCGGTGTAGCCCGCGGCGAGGGCGACACTGTGGGGGTGGGGCGTTGACGT 73625
QY 1477 -----CAGAAAGCCGTG 1488
Db 73626 CTCCCGTGGCATGCGCCAGCGAGCAGAGCCAGACCCAGGCTGTGCGCAAGCCCGTG 73685
QY 1489 TCCCGGTGCTGCGGCGAGTGTGCGAGGAGGCGCAGGTGTGCGCGGGTCAAGGGGTTCACATCG 1548
Db 73686 TCCCGGTGCTGCGGCGAGTGTGCGAGGAGGCGCAGGTGTGCGCGGGTCAAGGGGTTCACATCG 73745
QY 1549 TCGTGTACAGTGTGTGACTGTGAGGCGGCGAGCTACCGGCAAAAC----- 1597
Db 73746 TCGTGTACAGTGTGTGACTGTGAGGCGGCGAGCTACCGGCAAAACCGAGTGAAGCG 73805
QY 1598 ----- 1597
Db 73806 CCTTCCGCGAGCGGGGTGGGAACGACAGAGGAGGTCTGTCCAAAGTCTGTGACTGT 73865
QY 1598 -----CAGAC 1602
Db 73866 GAGACCAAGCCCAAGGGGAGCAAGACAAACCCAGCCCTCTCTCTCTCAAGAC 73925
QY 1603 GACATGCGCTGCACCTTTGTGTGCGAGATGATGTGTCCCGGAGCAACACAGCTGC 1662
Db 73926 GACATGCGCTGCACCTTTGTGTGCGAGATGATGTGTCCCGGAGCAACACAGCTGC 73985
QY 1663 TTCCCGCGAGGCTGTGCTGTGCGAGTGGGAGACCGGCTGTGTGCTGTGCTGTGCTGTG 1722
Db 73986 TTCCCGCGAGGCTGTGCTGTGCGAGTGGGAGACCGGCTGTGTGCTGTGCTGTGCTGTG 74045
QY 1723 CTGTGAGGCTGTGCGGCTTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1782
Db 74046 CTGTGAGGCTGTGCGGCTTGTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 74105
QY 1783 GACAGCCCACTGTGAGGCTGTGAGGCGGCGCGCTGTGCTGTGCTGTGCTGTGCTGTG 1842
Db 74106 GACAGCCCACTGTGAGGCTGTGAGGCGGCGCGCTGTGCTGTGCTGTGCTGTGCTGTG 74165
QY 1843 GGCCTGTGTGCTGTGAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1902
Db 74166 GGCCTGTGTGCTGTGAGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 74225
QY 1903 GCCCAGCAGCCCTGTGTCCACCTCCGCTCAAGGCGCTGTGCTGTGAGCAGCTTCTGTGAG 1962
Db 74226 GCCCAGCAGCCCTGTGTCCACCTCCGCTCAAGGCGCTGTGCTGTGAGCAGCTTCTGTGAG 74285
QY 1963 GCGGCGCAATCTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2022
Db 74286 GCGGCGCAATCTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 74345
QY 2023 TGCCTGTGCGGCGCGCTGTGCGCTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCG 2082
Db 74346 TGCCTGTGCGGCGCGCTGTGCGCTGTGCGTGTGCGTGTGCGTGTGCGTGTGCGTGTGCG 74405
QY 2083 CTGTGCACTGTGTACTGTGTGCGCTTCCGCGGAGTGTGAGCGAGTGTGAGCGACATGTG 2142
Db 74406 CTGTGCACTGTGTACTGTGTGCGCTTCCGCGGAGTGTGAGCGAGTGTGAGCGACATGTG 74465
QY 2143 CCGAGGAGGCGCTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2202
Db 74466 CCGAGGAGGCGCTGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 74525

```

Query Match	Best Local Similarity	Matches 1788; Conservative	Score 1574; DB 6; Length 1788;	Pred. No. 4.2e-201; Mismatches 0; Indels 204; Gaps 1;
QY 2203	CGCACCAATGGCAGCGTGGCCCTTTCGTCGCTTCCTGGGAGACTTCTCGGTGGGAGCCAG	2262		
Db 74526	GCCACCAATGCCAGCGTGGCCCTTTCGTCGCTTCCTGGGAGACTTCTCGGTGGGAGCCAG	74585		
QY 2263	CCGGGCTGCTACACCGTCCGCCGTGGCCCTCACCTTGGCCATCTGGCCCTACTTCATCAC	2322		
Db 74586	CCGGGCTGCTACACCGTCCGCCGTGGCCCTCACCTTGGCCATCTGGCCCTACTTCATCAC	74645		
QY 2323	TGGGCTCTCTTGTGTGTCGCCCTCTCTGGCCAAATGTGCAAGTGTCTCTCAAGCGCCGCTGCGAG	2382		
Db 74646	TGGGCTCTCTTGTGTGTCGCCCTCTCTGGCCAAATGTGCAAGTGTCTCTCAAGCGCCGCTGCGAG	74705		
QY 2383	ATGGGCGCCCTCTCTCTCTCTGTCTCTGGGCAATCCGAGCGCTTCACCTGCGCCAGTGT	2442		
Db 74706	ATGGGCGCCCTCTCTCTCTCTGTCTCTGGGCAATCCGAGCGCTTCACCTGCGCCAGTGT	74765		
QY 2443	TACCTGCTCATGCGGCGAGCCAGGCTCAACACCCCGAGTCTCTCTGGAGGGGGCCCT	2502		
Db 74766	TACCTGCTCATGCGGCGAGCCAGGCTCAACACCCCGAGTCTCTCTGGAGGGGGCCCT	74825		
QY 2503	GGGGATGCCAGGCGAGATATACGGGAAACAGAGAAATCAGGGAGAAATATAGTGA	2559		
Db 74826	GGGGATGCCAGGCGAGATATACGGGAAACAGAGAAATCAGGGAGAAATATAGTGA	74882		
RESULT 11				
LOCUS	AX451560	1788 bp	DNA	linear
DEFINITION	Sequence 4 from Patent WO0224885.			
ACCESSION	AX451560			
VERSION	AX451560.1	GI:21698546		
KEYWORDS				
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
TITLE	1			
JOURNAL	Kossida, S.			
FEATURES	Regulation of human g protein-coupled receptor			
source	Patent: WO 0224885-A 4 28-MAR-2002;			
	Bayer Aktiengesellschaft (DE)			
	Location/Qualifiers			
	1. 1788			
	/organism="Homo sapiens"			
	/db_xref="taxon:9606"			
BASE COUNT	236 a 620 c 593 g 339 t			
ORIGIN				

OY	790	AACAGAGCAGCGTGCAGGGTGATGCTGTGCTTTGGCTCCGTGACGCCCAAGCCCTTC	849
Db	301	AACAGAGCAGCGTGCAGGGTGATGCTGTGCTTTGGCTCCGTGACGCCCAAGCCCTTC	360
OY	850	TTCAACTACAGCATCACACAGGCTCTCGCCCCAAGGTGTGGGTGGACCAGCAGGACTCTGG	909
Db	361	TTCAACTACAGCATCACACAGGCTCTCGCCCCAAGGTGTGGGTGGACCAGCAGGACTCTGG	420
OY	910	CTGACCTCTGACCTGTGTATGGGGCTTGCCGGCATGGCCAAATGGGCACAGGCTCTTGGC	969
Db	421	CTGACCTCTGACCTGTGTATGGGGCTTGCCGGCATGGCCAAATGGGCACAGGCTCTTGGC	480
OY	970	TTCTCCAGAGGGGTGGCCAGCTGCAGAGATTCCCAGCTACGTGAAGTCAGCACTGGCC	1029
Db	481	TTCTCCAGAGGGGTGGCCAGCTGCAGAGATTCCCAGCTACGTGAAGTCAGCACTGGCC	540
OY	1030	CTGGCCACCAACCCCGGCTTCTGCTCTGGCCCTGGCCGAGAAGGAGACAGAGTCTGGAGAG	1089
Db	541	CTGGCCACCAACCCCGGCTTCTGCTCTGGCCCTGGCCGAGAAGGAGACAGAGTCTGGAGAG	600
OY	1090	GACGTGGTGGGCCAGCGCTGCCCGCAGGTGTGACTGTGCATCAGCTGCAGAACGTGAGCGCA	1149
Db	601	GACGTGGTGGGCCAGCGCTGCCCGCAGGTGTGACTGTGCATCAGCTGCAGAACGTGAGCGCA	660
OY	1150	GGGCTTAATCACCAACAGAGCTTCTGTCTATACGACAGCTGTATATAGCTGTGGCCCAAGCC	1209
Db	661	GGGCTTAATCACCAACAGAGCTTCTGTCTATACGACAGCTGTATATAGCTGTGGCCCAAGCC	720
OY	1210	CTGCACAACACTCTTCAGTCAACAGCCTCAGGTGGCCCCGCGCAGAGACCCCGTGAAAGCC	1269
Db	721	CTGCACAACACTCTTCAGTCAACAGCCTCAGGTGGCCCCGCGCAGAGACCCCGTGAAAGCC	780
OY	1270	TGCGAGGCTCTGGAGAACATGTACAACTGTGACCTTCACAGTGGGGGCGCTGCCTGTGGGG 	1329
Db	781	TGG-----	783
OY	1330	TTGCACACAGCAGCGGAACCTGTGACATGTGATACGACTTAAGCTGTGGGTGTGGCAGGGC	1389
Db	784	-----	783
OY	1390	TCATGTGCCACAGGCTCYCACGACGTGGGAGTTCAACGCGAGCCTCAGACAGAGCGCCTG	1449
Db	784	-----	783
OY	1450	AAGATCCGCTGGGCACACGTCTGAACACAGAGGCCGTGTCCCGTGTCTCGCGCAGTGC 	1509
Db	784	-----CAGAACCCGTTCCCGGTGCCTCGCGCAGTGC	816
OY	1510	CAGAGAGGGCCAGAGTGCGCCCGGGTCAAGGGTTCCATCTCTGTGTACAGACTGTGTGAC	1569
Db	817	CAGAGAGGGCCAGAGTGCGCCCGGGTCAAGGGTTCCATCTCTGTGTACAGACTGTGTGAC	876
OY	1570	TGCGAGCGGGGACAGCTACCGGCAAAAACCAAGACGACATGCGCTGCACCTTTGTGGCCAG	1629
Db	877	TGCGAGCGGGGACAGCTACCGGCAAAAACCAAGACGACATGCGCTGCACCTTTGTGGCCAG	936
OY	1630	GATGAGTGTATCCCGGAGCAGAACACAGCTGTCTTCCGCCAGAGGTCTCGATTCTTGCA	1689
Db	937	GATGAGTGTATCCCGGAGCAGAACACAGCTGTCTTCCGCCAGAGGTCTCGATTCTTGCA	996
OY	1690	TGGGGCAGACCGGCTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTG	1749
Db	997	TGGGGCAGACCGGCTGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTG	1056
OY	1750	CTGGCTCTTTGGGGGCTGTCTGTCAACATCCGAGACGCCCACTGGTTCAAGGCTTCGGGG	1809
Db	1057	CTGGCTCTTTGGGGGCTGTCTGTCAACATCCGAGACGCCCACTGGTTCAAGGCTTCGGGG	1116
OY	1810	GGGCGCCCTGGCTGTCTTTGGCTGTGTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	1869
Db	1117	GGGCGCCCTGGCTGTCTTTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	1176
OY	1870	TTTCCTTGGCCAGCCACCCCTGTGCCGATGCTGGCCCAAGACCCCTTGTCCCACTCTCCG	1929

Db 541 CCATCTTCTCCGACAGTGCACAGTGCAGCCGGGTGACGTGACAGCCGTTGTGACACTG 600
 QY 601 CTGAGAGAGTTCGGCTGGAAGTGGTGGCCGCTGAGGACAGACAGATATGAGCCGG 660
 Db 601 TTGCAATTTTCACTGGAAGTGGTGGCTGCTTGGTATGATGATGATGATGATGATG 660
 QY 661 CAGGAGCCAGACATCTTCTCGCCCTGGCCGCGACGCGGACATGTCATGCGCGACAG 720
 Db 661 GAAGGTCTGACATCTTCTTGTGTGTGCGCAACTACAGAGGTATCTGATTCACACAG 720
 QY 721 GGCCTGTGCGCGTCCCGCCGCGATGATGCTGCGGCTGGGGAGGTGACAGAGCTCTG 780
 Db 721 GGCCTGTGCGCGTCCCGCCGCGATGATGCTGCGGCTGGGGAGGTGATGATGATG 780
 QY 781 CACAGAGTGAACAGAGAGGCTGAGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 Db 781 CGCCAGTGAACAGAGAGGCTGAGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
 QY 841 CAGGAGCCAGACATCTTCACTAGACATGACAGAGGCTCTGCGCCAGAGTGTGGTGGCCAGC 900
 Db 841 TACTCCCTTTTACCTACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 900
 QY 901 GAGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 960
 Db 901 GAGTCTGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 960
 QY 961 GTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1020
 Db 961 GTTCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1020
 QY 1021 CACCTGGGCTGGGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1080
 Db 1021 CGCCTGGGCTGGGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1080
 QY 1081 CTGAG 1140
 Db 1081 CTGAG 1140
 QY 1141 GTGAGCGAG 1176
 Db 1138 CTGATCATCTGGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1197
 QY 1177 GTCTACGAGAGTGTGTATAG 1236
 Db 1198 ACCTATGAGAGTGTGTATAG 1257
 QY 1237 TCAG 1296
 Db 1258 TCAGATGAG 1317
 QY 1297 CTGACCTTCACTGAG 1356
 Db 1318 ATGATGATGAG 1377
 QY 1357 GAGTACAG 1416
 Db 1378 GAAATATGAG 1437
 QY 1417 AGTGTCAAG 1476
 Db 1438 ACCTTCAAG 1491
 QY 1477 CAGAAGCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCC 1536
 Db 1492 CAGTGTGAGAGTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCC 1551
 QY 1537 GAGTGTGAGAGTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCC 1596
 Db 1552 GAGTGTGAGAGTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCC 1611
 QY 1597 CAGAAGCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCC 1656
 Db 1612 CAGAAGCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCCGCTGTCC 1671

QY 1657 CGCTGCTTCGCGCGAGAGTCTCGGTTCTGTCATGAGGAGCGAGCGGCTGTGCTGTGCTG 1716
 Db 1672 ACCTGCTTACTCTGCGAGGCTCCAGATTTCTGCTGTGGGAGCGACAGCTGTGCTGCTGCTG 1731
 QY 1717 CTCTGCTGTGAGACCTGAGCGCTGGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 1776
 Db 1732 CTCTGCTGTGAGACCTGAGCGCTGGGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 1791
 QY 1777 CATGGGAGACCCACATGCTGAGGCTCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1836
 Db 1792 TACTGGGAGACCCCTGTTGTCAGGCGCTGACAGTGGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1851
 QY 1837 TGCCCTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1896
 Db 1852 TGCCCTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1911
 QY 1897 TGCCCTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1956
 Db 1912 TGCCCTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1971
 QY 1957 CTGAGAGGCGCGAGATTTGTGTGAGTCAAACTGCTCTGAGCTGTGAGCTGGGAGACCGGCTG 2016
 Db 1972 CTGAGAGGCGCGAGATTTGTGTGAGTCAAACTGCTCTGAGCTGTGAGCTGGGAGACCGGCTG 2031
 QY 2017 AGTGGCTGCTGCGGGGGGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 2076
 Db 2032 TGACAGCTTACTCTGCGGGGCGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTG 2091
 QY 2077 GTGCACTGTGACCTGTGTACTGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2136
 Db 2092 GCTGCACTGTGACCTGTGTACTGTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2151
 QY 2137 ATGCTGCCAGAGAGAGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2196
 Db 2152 GTGCTGCCAGAGAGAGGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2211
 QY 2197 GCGCAGCGCAACAAATGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2256
 Db 2212 GTGCAATCAACAAATGCGACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2271
 QY 2257 AGCCAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2316
 Db 2272 AGCCAGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2331
 QY 2317 ATCACTG 2376
 Db 2332 ATCACTG 2391
 QY 2377 GTGCAATGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2436
 Db 2392 GTGCAATGAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2451
 QY 2437 AGTGTACTG 2496
 Db 2452 AAATGTACTG 2511
 QY 2497 GGCCTGGGAGTGGCCCAAGGCGCAATGACGGGA 2530
 Db 2512 AGCCCAAGGAGATCATGATGGGAATGTGTGTA 2545

RESULT 13
 LOCUS AY032620 2577 bp mRNA linear ROD 05-FEB-2002
 DEFINITION Rattus norvegicus sweet taste receptor T1R3 mRNA, complete cds.
 ACCESSION AY032620
 VERSION AY032620.1 GI:15147674
 KEYWORDS
 SOURCE Rattus norvegicus.
 ORGANISM Rattus norvegicus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

Db	1438	ACCTTCAACGGCACACCTCTCAGCTGCGACACTCGAATAATGATTGGC-----CAGGCAC	1491
Qy	1477	CAGAAAGCCCGTGTCCGGGTCTGCGGCACTGCGACGAGAGGCCAGGTGCGCGGTCAAG	1536
Db	1492	CAGGTGCACAGCTCCACAGTGTCTCCGGCAGATGCAAAAGATGGCCAGGTGCGCAGACTAAG	1551
Qy	1537	GGGTTCACATCTCTGCTCACTGACGATGATGTGACATGTGGAGAGGCGGACGATACCGGGCAAAAC	1596
Db	1552	GACTTTCATTCTCTCTGTATGACTGTGTGTGACTGCAAGGCAAGGAGACTACCGGAAGCAT	1611
Qy	1597	CCAGACGACATCGCCTCGACCTCTTGTGTGGCAGAGATGAGTGTCCCGGAGGCAAGACAA	1656
Db	1612	CCAGATGACCTTACACTGTACTTCATGAGAGAGATCAGTGGTCCCAAGAAAAAGACAA	1671
Qy	1657	CGCTGCTTCCCGCCGAGGTCTCGGTTCTCTGGCATGGGGCGAGCCGAGCTGTGCTGCTG	1716
Db	1672	ACCTGTCTTACCTCCACAGGCCCAATTTTCTGGCTGGGGGAGGACACCTGTGCTGCACATT	1731
Qy	1717	CTCCGTGCTGTGACCTCGAGCGCGGGGCGCTTGTCTGTGGCTGTGTTTCTTCATC	1776
Db	1732	CTCCGTGCTGTGCTCGAGTGTGTGGGCTTACACTGTGCTGCCCTGTGGGCTCTTTGTCCAC	1791
Qy	1777	CATCGGACACGCCCACTGCTTCAAGGCTCGGGGGGAGCCCTGTGCTGTGGCTGTG	1836
Db	1792	TACTGTGGACACCCCTCTTGTTCACAGGCTTCAAGGTGTCACCTGTGTTTGCCCTGATC	1851
Qy	1837	TGCCGTGGGCTGTGCTGCTCAGCGTCTCTGTGTCACCGGCGGACGCCCTGCCGA	1896
Db	1852	TGCCCTAGGCTCTTCTGCTCAGTGTCTTCTTCCACAGACGACACACGCTGTGCAAC	1911
Qy	1897	TGCCGTGGGCGACGACCCCTGTGTCCACACTCCGCTCAGGGGCTGCTGTGACACTCTTC	1956
Db	1912	TGCCTTGGCCCAACCAACCAATGGCTCACCTCCCTCACAGGCTGCTGTGACACTCTTC	1971
Qy	1957	CTGCGAGGGGGCGAGATCTTCGTGAGATCAGAACTGCTGTAGCTGGGCGAAGCCGGCTG	2016
Db	1972	CTGCGAGAGCGCGAATCTTGTGTGAGTCTGTGACTGTGCATGTGTTGGGCAAACTGGCT	2031
Qy	2017	AGTGGCTGCTGCGGGGGCCCTGGGCGCTGGCTGTGTGTGCTGTGCGCAATGCTGTGAG	2076
Db	2032	TGCACACTACCTTCGGGGGCCCCGTGGCTGTGGTGTGTACTGCTGTGCGCACTTGTGGAG	2091
Qy	2077	GTGCGACTGTGCACTGTGTACTGTGTGTGTGGCTTCCCGCGGAGGTGTGTACAGACTGGCAC	2136
Db	2092	GCTGCACTATGTGCTGTGTACTGTGTATGTGGCTTCCCTCCAGAGGTGTGTGACAATTTGGCAG	2151
Qy	2137	ATGCTGGCCACGAGAGCGGCTGTGTGCACTGTGCGGACAGCGCTCGTGGGTGACCTTGGCCTA	2196
Db	2152	GTGCTGGCCACGAGAGTACTGTGAACAATGCGCGCATGCTTCCGTGGGTACGCTTGGGCTTG	2211
Qy	2197	GCGCAGCGCCACAATGTCCAGCGCTGGCTTCTGTGCTTCTGTGGGCACTTTCCTGTGGTGG	2256
Db	2212	GTGCACTACACCAATGTGCAAGTGTATAGCTTTCCTGTGCTTCTGTGGGCACTTTCCTGTGAACG	2271
Qy	2257	AGCCAGCGCGGCTCTCTCAACCGGCGGCCGCGGCTCACTGCTTGGCCATGTGCTGGCTACTTTC	2316
Db	2272	AGCCAGCGCTGTGCTCTTAACCGGTGCTGCGCGGCTTACCTTGGCCATGTGTACTTAATTC	2331
Qy	2317	ATCACCTGGGTCTCTTGTGTGCCCTCTGGCCCAATGTGCAAGTGTCTCTCAAGGCCGCG	2376
Db	2332	ATCACTTGGTGTCTTGTGTGGCCCTCTGCTGCTAATGTGAGGTGTGCTTACACGACAGT	2391
Qy	2377	GTGCAAGATGGGCGGCTCTGTGCTGTGTCTGTGGCATCTGTGGCTGCTTCCACTGCCC	2436
Db	2392	GTGCAAGATGGGTGTACTTAATCTTGTGTGCGCTGTGGGCACTGTGCGCACTTCCACTGTGCC	2451
Qy	2437	AGGTGTTCACGTGCATGCGGCGGACGCGGCTCAACACCCCGAGTTCCTCTGGGAGGG	2496
Db	2452	AAATGTCTATGTACTTCTGTGTGCTGTGCAAGACTTAACACCCAGAGTCTTCTCTGGGAGGG	2511
Qy	2497	GGCCCTGGGATGTCCCAAGGCGCAATGACGGGA	2530

[illegible]

[illegible]

Db	Accession	Source	Organism	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	REFERENCE
Db	1672	GCCTGCTTACCTGCGAGGCCCAAGTTCTTGAGTTTGCGGGAGACAGTTGTGCTGAC	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Qy	1717	CTCCTGCTGCTGAGGCTGAGCGCTGGAGCCTTGCTGAGCTGCTTTGGGGCTGTGCTTAC	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Db	1732	CTTCCTGCTGCTTGGCTGGTGTGGGTGTAGCACTGCTGCTCTGGGGCTCTCTCTCCAC	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Qy	1777	CATCGGAGACGCCACTGTGTAGGCGTCMGGGGGGGGCCCTTGCGCTTGCTTGGCTGGTG	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Db	1792	CAGTGGAGAGGCCCTCTTGTCCAGGCTCGAGGTGGCTACACTTGTGCTTGGCTGTATC	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Qy	1837	TGCCTGAGGCTGTGTGCTCAGCTCCCTCTGTTTCCCTGGGACGCCAGCCCTGCGCA	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Db	1852	TGCCTAGGCTCTTGTGCTCAGTGTCTCTGTGTTCCAGGGCGGCCAAGCTCTCCAGC	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Qy	1897	TGCTGAGGCCAGACGCCCTTGTCACACTCCCGCTCAGAGGGGTGCTTACGACACTTTC	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Db	1912	TGCCTTGGCAACAACCAATGGCTCAGCTCCCTCCACAGGCTGCTTACGACACTCTTC	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Qy	1957	CTGCAAGCGGGCGAGATCTTGTGGAGTCAGAACTCCCTAGCTGGGAGACAGCGGTG	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Db	1972	CTGCAAGCACTGACACTTCTTGTGAGTCTAGAGCTGACACTGAGCTGGGCAACTGGCTA	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Qy	2017	AGTGGCTGCTCTCGGGGGGCCCTGGGCTGAGCTGATGATGCTGCTGCGCACTGTGTGAG	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Db	2032	TGCAGTACTCTTGGGGAGCTGTGGGCTGAGTGTGTTGTTGGTGGTCCACTTTTGTGAG	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Qy	2077	GTGCGACTGTGACCTGTGACTGTGTGGCTTCCCGCGGAGTGTGTACGAGACTGGCAG	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Db	2092	GGAGCACTATGTGCTGTGATTTGATTCGCTTCCACAGAGTGTGACACACTGTGTCA	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Qy	2137	ATGCTGCCACGAGAGCGCTGTGTCAGTCCGACACGCTCTCGGTGACGTTGCGCTTA	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Db	2152	GTGCTGCCACAGAGTACTGAGACACGACGACGTCGTCTCTGGTCAAGCTGGGCTTG	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Qy	2197	GGCGACGCCACCAATGACACGCTGGCTTCTCTGCTTCTGAGGCACTTCTGTGTGGG	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Db	2212	GTGCACTATACCAATGACATGTTACTCTTCTCTGCTTCTGAGGCACTTCTGTGTAG	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Qy	2257	AGCCAGCGCGGTGTGTACACAGTCCGCTGAGCTCAGCTTGTGCGCATGTGAGCTACTTC	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Db	2272	AGCCAGCGCGGTGTGTACACAGTCCGCTGAGCTCAGCTTGTGCGCATGTGAGCTACTTC	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Qy	2317	ATCAGCTGGGTCTCTTGTGCTTCTGCGCATGTGAGAGTGTCTCAGGCCGCC	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Db	2332	ATCAGCTGGGTCTCTTGTGCTTCTGCGCATGTGAGAGTGTCTCAGGCCGCC	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Qy	2377	GTGCGAGTGGGGGCGCTCTGCTGTGTCGCGGAGCATGCTGAGCTTCCATCAGCTGCC	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Db	2392	GTGCGAGTGGGGGCGCTCTGCTGTGTCGCGGAGCATGCTGAGCTTCCATCAGCTGCC	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Qy	2437	AGGTGTTACCTGCTCATGCGGAGCGAGCGCTCAACACGCCGAGTCTTCTGTGGAGGG	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1r3)	AF311386	1	(bases 1 to 2577)	
Db	2452	AGGTGTTACCTGCTCATGCGGAGCGAGCGCTCAACACGCCGAGTCTTCTGTGGAGGG	Mus musculus	AF311386	mus musculus putative sweet taste receptor type 1 member 3 (Tas1				

AUTHORS Bachmanov, A.A., Li, X., Reed, D.R., Ohmen, J.D., Li, S., Chen, Z., Tordoff, M.G., deJong, P., Wu, C., West, D.B., Chatterjee, A., Ross, D.A. and Beauchamp, G.K.

TITLE Positional cloning of the mouse saccharin preference (Sac) locus

JOURNAL Chem. Senses 26 (7), 925-933 (2001)

MEDLINE 21439023

PubMed 11555487

REFERENCE 2 (bases 1 to 2577)

AUTHORS Li, X., Bachmanov, A.A., Li, S., Chen, Z., Tordoff, M.G., Beauchamp, G.K., deJong, P., Wu, C., Chen, L., West, D.B., Ross, D.A., Ohmen, J.D. and Reed, D.R.

TITLE Genetic, physical, and comparative map of the subtelomeric region of mouse Chromosome 4

JOURNAL Mamm. Genome 13 (1), 5-19 (2002)

MEDLINE 21635547

PubMed 11773963

REFERENCE 3 (bases 1 to 2577)

AUTHORS Bachmanov, A.A., Li, X., Reed, D.R., Ohmen, J.D., Li, S., Chen, Z., Tordoff, M.G., deJong, P., Wu, C., West, D.B., Chatterjee, A., Ross, D.A. and Beauchamp, G.K.

TITLE Direct Submission

JOURNAL Submitted (05-OCT-2000) Monell Chemical Senses Center, 3500 Market Street, Philadelphia, PA 19104, USA

FEATURES

source

1..2577

/organism="Mus musculus"

/strain="C57BL/6J"

/db_xref="taxon:10090"

/chromosome="4"

/map="83.0CM"

1..2577

/gene="Tas1r3"

/allele="Sac-b"

1..2577

/gene="Tas1r3"

/codon_start=1

/product="putative sweet taste receptor type 1 member 3"

/protein_id="AL08425.1"

/db_xref="GI:15808695"

/translation="MPALIMGLSLAFLFELGKASISLCSQFKAGDYLILGLFPIG STEFATLNRTOPISTPCNRFSPILGLFLAMKMAVEELINNSALLGLRLGLDPT CSRPVYTKMSLMFLAKVSQSIAATCNTQYQPRVLAIVGRISSELAITLTKFSEF LMPQVYSASMDRLSDRETFPSDFRTVPSDRVOLAIVLILSNMVAALSDSDY GREGISLFSLSANARICIAHEGLVPQHTSGOGLKVLADVLQNVQSKVQVVLAS ARAVYSLFYSIHGSLSPKVMYSESLSDILVMTLPNLAIVTVGLVGLALDEF SHVETHLALADPARFCASINAELEBEHMGRCRCDIMQNTSSGLIQLVLSGO LHHQIFATYAAVYSAVALHNTLQCNVSHCHVSEHLIPMLLENMTMSPHADLTQ FPAEGVNDMEYDLKMWKOSPPTVLTHTVGTFFNCTILOOSKMTPEGNQVPVSCSRQC KDQVRRVGFHSCCYDCVCKAGSYRKHDDCTPCNODWSPEKSTACLERRPF LAMGEVNVLSLLILCLVGLAALAGLSVHHMDSPLVQASGSQCFGLICGLFCL SVLFFGRSSASCAQOPMAHLPLTGCLSTFLQAEFESELEPLSMNMVLSYLR GLWAMLVLLATVEAALCAWLIAPPEVTVGLVPEVLEHCHVRSVSLGVLK TNAMLAFELFELFELVOSQGRYRNARGLTFMALVFTVWSFPLIANOVYQYAV QKGAIIYVCAIGILVTHPRKCYLLMLPLINTDFELGRNAKAADENSQGGNAQGH NE"

BASE COUNT 519 a 735 c 689 g 634 t

ORIGIN

Query Match 56.9%; Score 1454.8; DB 10; Length 2577;

Best Local Similarity 74.3%; Pred. No. 3.2e-185;

Matches 1891; Conservative 0; Mismatches 622; Indels 33; Gaps 3;

181 GTGTGACACAGTTCCTCTCAAAAGCCCTGCTGTGGACATGGCCATGAATAATGGCCGTG 240

181 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

181 CGGTGACACAGTTCCTCAAAAGCCCTGCTGTGGACATGGCCATGAATAATGGCCGTG 240

241 GAGAGATCAACAACAAGTGGATCTGCTGCGCGGCTGCGCTGGGCTACACCTCTT 300

241 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

241 GAGAGATCAACAACAAGTGGATCTGCTGCGCGGCTGCGCTGGGCTACACCTCTT 300

301 GATACGCTGCGGAGGCTGTGGTGGCCATGAGCCAGCCAGCTCATGTTCTGGCCAAAGCA 360

301 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

301 GATACGCTGCGGAGGCTGTGGTGGCCATGAGCCAGCCAGCTCATGTTCTGGCCAAAGCA 360

361 GGCAGCGCGACATCGCGCGCTACTGCACTACACGACGATACACGCGCGCTGTGCTGCT 420

361 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

361 GGCAGCGCGACATCGCGCGCTACTGCACTACACGACGATACACGCGCGCTGTGCTGCT 420

421 GTGATGCGGCG 480

421 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

421 GTGATGCGGCG 480

481 GTGATGCGGCG 540

481 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

481 GTGATGCGGCG 540

541 CCCTCTCTTTCGCGACCGTGCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600

541 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

541 CCCTCTCTTTCGCGACCGTGCACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 600

601 CTGACGAGATGTCG 660

601 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

601 CTGACGAGATGTCG 660

661 CAGGCGCTGACACATCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720

661 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

661 CAGGCGCTGACACATCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 720

721 GAGGCTGACACATCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780

721 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

721 GAGGCTGACACATCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 780

781 CACGAGTGAACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 840

781 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

781 CACGAGTGAACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 840

841 CAGGCGCTGACACATCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900

841 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

841 CAGGCGCTGACACATCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 900

901 GAGGCGCTGACACATCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960

901 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

901 GAGGCGCTGACACATCTCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 960

961 GTGCTGCGCTCTCCGACGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020

961 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

961 GTGCTGCGCTCTCCGACGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1020

961 GTGCTGCGCTCTCCGACGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080

961 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

961 GTGCTGCGCTCTCCGACGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1080

1021 CACCTGCGCTGCG 1132

1021 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

1021 CACCTGCGCTGCG 1132

1078 CTGAGGAGACATGATGATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1137

1078 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

1078 CTGAGGAGACATGATGATGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1137

1133 -----TGCAAGACGTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1176

1133 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

1133 -----TGCAAGACGTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1176

1138 CTATCATCTGCGCTGTGCG 1197

1138 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

1138 CTATCATCTGCGCTGTGCG 1197

1177 GTCTACGACGCTGTGTATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1236

1177 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

1177 GTCTACGACGCTGTGTATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1236

1198 ACCTATGACGCTGTGTATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1257

1198 | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

1198 ACCTATGACGCTGTGTATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1257

OY	1237	TCAGAGCTCCCCGGCAGAGACCCCGTGAAGACCCCTGGCAGACTCTGTGAGAACATGTACAC	1236
Db	1258	TCACATTGCGACGATATCAGACACTGTTTCTTACCTGGCAGCTCTGTGAGAACATGTACAT	1317
OY	1297	CTGACCTTCACGCTGGGGCGGGCTGCCGCTGCGGGTTTCGACAGCAGCGGAACAGTGGACATG	1356
Db	1318	ATGAGTTTCCATGTGTGACACTTGTGACACTACAGTTTGGATGCTGAAGGGAATGTAGACATG	1377
OY	1357	GAGTACGACCTGAAGCTGTGTGGTGTGGCAGAGGCTCAGTGTCCAGGCTCCACGAGTGGGC	1416
Db	1378	GAATATGACCTGAAGATGTGGTGTGGCAGAGGCCCTACACCTGTATTACATACTGTGGC	1437
OY	1417	AGGTTCAAGGGGACGCCACAGACAGAGCGGCTGAAGATCCGTGGCACAACGCTGCACAC	1476
Db	1438	ACCTTCACGGCACCCCTTACGCTGACGACAGCTTAAATATTACTGGC-----CAGCCAC	1491
OY	1477	CAGAAAGCCGCTGTCCGCGTCTGCGCGGACAGTGCACAGAGAGGGCGCAGGTGCGCGGGTCAAG	1536
Db	1492	CAGGTGGCAGCTGTCCGAGTGTTCGCCGACGTGAAGATGGCGCAGTGTCCCGGATTAAG	1551
OY	1537	GGGTTCCACTCTCTGTCTGTACGACTGTGTGGACTGTGGAGGGCGGAGCTACCGGCAAAAC	1596
Db	1552	GGCTTTATTCCTGTGCTCTGTATGACTGCTGTGAGACTGCAGAGGGCGGACGCTACCGGAAGCAT	1611
OY	1597	CCAAACACACATCCGCTGACACTTTTGTGGCGACAGATAGATGGTCCCGGAGGGAAGCACA	1656
Db	1612	CCAAATATCTTCACTCTGTACTCTCATATTAATTAATTAATTAATTAATTAATTAATTAATTAAT	1671
OY	1657	CGCTGCTTCCGCGCAGAGTCTCGGTTCTGTGACATGGGGGGGACCCGCGCTGTGCTGTGATG	1716
Db	1672	GCCTGCTTACCTGCGAGGCCCAAGTTTCTGTGCTGTGGGGGGAACAGATTGTGCTGTACCTC	1731
OY	1717	CTCTGCTGCTGAAGCTGTGGCGGTGGGCTTGTGTGCTGTGCTTGTGGGCTGTTCGTTTAC	1776
Db	1732	CTCTGCTGCTTGTGCTGTGCTGTGGGTTCAAGCACTGGCTGTCTGGGGCTCTCTGTCCAC	1791
OY	1777	CATGGGGACAGCCACAGGTTTCAAGGCTGTGGGGGGGCCCCCTGTGCTTGTGGCTGGTG	1836
Db	1792	CAGTGGACAGCCCTTGTGTGCCGCTTCAAGGTGAGGTCTCACTTGTGCTTGTGGCTGTATC	1851
OY	1837	TGCTGTGGGCTGTGCTGTGCTCAGCGTCTCTCTCTGTTCCCTGTGGCGACCCAGCCCTCCCGGA	1896
Db	1852	TGCTTAGGCTCTTGTGTGCTCAGTGTCTTCTTCTTGTTCGCCAGGGCGGCCAAGCTGTGTCCAGC	1911
OY	1897	TGCTGTGGCCACAGCCCTTGTGCCACCTTCGCGCTACGCGGCTGTGCTGTAGCACACTTTC	1956
Db	1912	TGCTTGTGACAAACACCAATGGCTCACCTCCCTCTCACAGGCTGTGCTGTAGCACACTTTC	1971
OY	1957	CTGACAGCGGCGCCGAGATCTTCTGTGTAGTACAGAACTGCTCTGTAGCTGGGCAAGCCGCTG	2016
Db	1972	CTGTACACACTGTAGACTTGTGTGTGAGTGTGTGAGCTGTGCCACTAGCTGGGGCAAACTGTGCTA	2031
OY	2017	AGTGGTGTGCGGGGGGGGCCCTGGGGCTGGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG	2076
Db	2032	TGCAAGCTTACTTGTGGGAGCTCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGGCTGTGG	2091
OY	2077	GTCGCACTGTGACCTGTGTACTGTGTGTGGCTTCCGCCCGAGAGTGTGTAGCGAGCTGAGAC	2136
Db	2092	GCACACATATGTGTGCTGTATTTGATATCGCTTTCGCCACAGAGGTGTGTAGACATGTGCTGCA	2151
OY	2137	ATGCTGCCCAAGAGCGGCTGTGTGACCTGTGCGCACACGCTCCTGTGGGTACGTTTGGGCTTA	2196
Db	2152	GTCGTGTGCCACAGAGGTACTG	2211
OY	2197	GGCAGCGCACCAATGTGCAGGTGTGGGCTTGTCTGTCTGTGGGCACTTGTCTGTGGTAAAG	2256
Db	2212	GTCACATACACATGTGAATGTTAAGCTTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTG	2271
OY	2257	AGCCAGCGGAGCTGTACAAACCGTGGCCGTGGGCTTACCTTGTGCCATGTGTGGCTACTTTC	2316
Db	2272	AGCCAGCGTGGCGCTGTACAAACGCTGGCCGTGGGCTTACCTTGTGCCATGTGTACTTATTTTC	2331
OY	2317	ATCACCTGTGGGCTCTCTTGTGGCCCTTGTGGCCATGTGTGACAGTGTGTCTCAGGCCGCC	2376

Accession	Sequence	Position
Dd	ATCACTGGGCTCTTTTGAGCCCTCTGCAATGTCAAGTGGCCCTACCAAGCACT	2391
Dd	2332 ATCACTGGGCTCTTTTGAGCCCTCTGCAATGTCAAGTGGCCCTACCAAGCACT	2391
Qy	2377 GTGAGATGGGGCGGCGCCCTGCTGCTGTCGTGTCGGGATCTGGTGCCTCACTGCC	2436
Dd	2332 GTGAGATGGGCTGCTATCTAGTGTGTGCTGGCCCTGGGCACTTGCTCACTGCC	2451
Qy	2437 AGGTATTACCTGCTCATGCGGACCAAGGCTCAACACCCCGAGTCTTCTCGGAGAG	2496
Dd	2452 AAGTGTATGTGCTCTTTTGCTGGCCAAAGCTCAACACCCAGAGTTCTTCTGGGAAG	2511
Qy	2497 GGCCCTGGGATGCCCAAGCCAGAA	2522
Dd	2512 AATGCCAAGAAAGCAGATGAGAA	2537

Search completed: May 23, 2003, 16:46:17
Job time : 7172.01 secs

2317 ATCACCTGGGTCCTCTTGTGCCCCCTCTGGCCAATGTGCAGGTGCTCCTCAGGCCGCC 2376

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4-P5-4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 23, 2003, 11:20:10 ; Search time 534.875 Seconds
(without alignments)
10774.220 Million cell updates/sec

Title: US-09-927-315-14

Perfect score: 2559

Sequence: 1 atgtctggccctctgtctcct.....atcagggaacatcagtgta 2559

Scoring table: IDENTITY_NUC
Gapop 10.0 ; Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N.Geneseq.101002:*

1: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2557.4	99.9	2559	24	ABK47345
2	2557.4	99.9	3458	24	ABL38462
3	2555.8	99.9	2559	24	AAS97395
4	2555.8	99.9	3489	22	AAS14574
5	2537	99.1	2553	22	AAD17509
6	2514.4	98.3	2729	24	ABL55952
7	2090.6	81.7	2739	24	ABL16615
8	1947.4	76.1	8001	24	ABL38463
9	1897	74.1	3563	22	AAD17517

10	1666	65.1	2687	22	AAD17508	Human taste recept
11	1574	61.5	1788	24 <td>AAD34459</td> <td>Human GPCR CDNA #2</td>	AAD34459	Human GPCR CDNA #2
12	1470.8	57.5	2577	22 <td>AAD17514</td> <td>Rat taste receptor</td>	AAD17514	Rat taste receptor
13	1454.8	56.9	2577	24 <td>AAS97393</td> <td>Murine SACL gene C</td>	AAS97393	Murine SACL gene C
14	1083	42.3	1113	24 <td>AAD34458</td> <td>Human GPCR CDNA #1</td>	AAD34458	Human GPCR CDNA #1
15	871.6	34.1	11809	24 <td>AAS97394</td> <td>Murine SACL gene g</td>	AAS97394	Murine SACL gene g
16	848	33.1	1086	24 <td>AAS98148</td> <td>Human DNA for pote</td>	AAS98148	Human DNA for pote
17	404	15.8	405	22 <td>ABA08492</td> <td>Human taste recept</td>	ABA08492	Human taste recept
18	381	14.9	876	22 <td>AAD17507</td> <td>Human taste recept</td>	AAD17507	Human taste recept
19	378.8	14.8	558	22 <td>AAD17513</td> <td>Rat taste receptor</td>	AAD17513	Rat taste receptor
20	366	14.3	2526	22 <td>AAD19501</td> <td>Human novel G-prot</td>	AAD19501	Human novel G-prot
21	366	14.3	2951	22 <td>AAD19505</td> <td>Human novel G-prot</td>	AAD19505	Human novel G-prot
22	365.6	14.3	2759	24 <td>ABL55953</td> <td>Human G-protein co</td>	ABL55953	Human G-protein co
23	357	14.0	2292	22 <td>AAD19502</td> <td>Human novel G-prot</td>	AAD19502	Human novel G-prot
24	354.2	13.8	2526	22 <td>AAD17516</td> <td>Human taste recept</td>	AAD17516	Human taste recept
25	350.8	13.7	2771	21 <td>AAZ50743</td> <td>Rat sensory trans</td>	AAZ50743	Rat sensory trans
26	346.6	13.5	2333	21 <td>AAZ50745</td> <td>Human sensory tran</td>	AAZ50745	Human sensory tran
27	344.4	13.5	2579	21 <td>AAZ50744</td> <td>Mouse sensory tran</td>	AAZ50744	Mouse sensory tran
28	319.6	12.5	552	22 <td>AAD17512</td> <td>Mouse taste recept</td>	AAD17512	Mouse taste recept
29	284	11.1	2532	21 <td>AAZ38964</td> <td>Mouse GPCR-B4 poly</td>	AAZ38964	Mouse GPCR-B4 poly
30	279.4	10.9	2993	21 <td>AAZ58963</td> <td>Rat GPCR-B4 poly</td>	AAZ58963	Rat GPCR-B4 poly
31	218.6	8.5	2010	21 <td>AAZ58965</td> <td>Human GPCR-B4 poly</td>	AAZ58965	Human GPCR-B4 poly
32	173	6.8	2703	21 <td>AAZ50617</td> <td>Human calcium sens</td>	AAZ50617	Human calcium sens
33	173	6.8	2751	20 <td>AAZ31057</td> <td>Chimeric hCAR/hmgl</td>	AAZ31057	Chimeric hCAR/hmgl
34	173	6.8	2751	24 <td>AAI43276</td> <td>Chimeric human hCA</td>	AAI43276	Chimeric human hCA
35	173	6.8	2922	21 <td>AAZ50616</td> <td>Human calcium sens</td>	AAZ50616	Human calcium sens
36	173	6.8	2925	22 <td>AAC85786</td> <td>CDNA encoding chim</td>	AAC85786	CDNA encoding chim
37	173	6.8	3003	21 <td>AAZ50615</td> <td>Human calcium sens</td>	AAZ50615	Human calcium sens
38	173	6.8	3234	20 <td>AAZ90922</td> <td>Human calcium rece</td>	AAZ90922	Human calcium rece
39	173	6.8	3234	20 <td>AAZ31049</td> <td>Human CAR coding s</td>	AAZ31049	Human CAR coding s
40	173	6.8	3234	24 <td>AAI43298</td> <td>Human wild type ca</td>	AAI43298	Human wild type ca
41	173	6.8	3234	24 <td>AAI43299</td> <td>Human CAR CDNA seq</td>	AAI43299	Human CAR CDNA seq
42	173	6.8	3809	17 <td>AAI61382</td> <td>Parathyroid calculu</td>	AAI61382	Parathyroid calculu
43	173	6.8	3809	19 <td>AAV26964</td> <td>Human parathyroid</td>	AAV26964	Human parathyroid
44	173	6.8	3809	19 <td>AAV26964</td> <td>Human parathyroid</td>	AAV26964	Human parathyroid
45	173	6.8	3809	20 <td>AAZ25055</td> <td>Human parathyroid</td>	AAZ25055	Human parathyroid

ALIGNMENTS

RESULT 1
ABK47345
18-JUN-2002 (first entry)
CDNA sequence encoding human AXOR79.
AXOR79; human; seven transmembrane; G-protein coupled; receptor; ss;
gene; bacterial; viral; fungal; infection; protozoal; mantic depression;
human immunodeficiency virus; HIV; cancer; diabetes; anorexia; bulimia;
Parkinson's disease; heart failure; asthma; allergy; osteoporosis;
neurological disorder; schizophrenia; Huntington's disease.
Homo sapiens.
Key 1.2559 Location/Qualifiers
CDS /tag=a
/product="AXOR79 protein"

GB2364058-A.
16-JAN-2002.
03-MAY-2001; 2001GB-0010887.
05-MAY-2000; 2000US-0566161.
(SMIK) SMITHKLINE BEECHAM CORP.

QY 1561 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 1620
 DB 1561 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 1620
 QY 1621 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 1680
 DB 1621 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 1680
 QY 1681 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 1740
 DB 1681 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 1740
 QY 1741 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 1800
 DB 1741 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 1800
 QY 1801 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 1860
 DB 1801 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 1860
 QY 1861 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 1920
 DB 1861 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 1920
 QY 1921 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 1980
 DB 1921 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 1980
 QY 1981 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 2040
 DB 1981 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 2040
 QY 2041 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 2100
 DB 2041 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 2100
 QY 2101 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 2160
 DB 2101 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 2160
 QY 2161 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 2220
 DB 2161 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 2220
 QY 2221 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 2280
 DB 2221 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 2280
 QY 2281 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 2340
 DB 2281 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 2340
 QY 2341 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 2400
 DB 2341 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 2400
 QY 2401 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 2460
 DB 2401 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 2460
 QY 2461 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 2520
 DB 2461 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 2520
 QY 2521 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 2580
 DB 2521 TGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 2580

RESULT 2
 AAL38462
 ID AAL38462 standard; cdna: 3458 BP.
 XX

AC AAL38462;
 DT 15-AUG-2002 (first entry)
 XX
 DE cDNA encoding a human G-protein coupled receptor (GPCR).
 XX
 DE G-protein coupled receptor; GPCR; gene chip; human; immune response;
 KW chromosomal position; transgenic animal; gene therapy; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 30..2588
 FT /tag= a
 FT /product= "Human G-protein coupled receptor"
 XX
 XX WO200230981-A1.
 XX
 XX 18-APR-2002.
 XX
 XX 13-MAR-2001; 2001WO-US07832.
 XX
 XX 10-OCT-2000; 2000US-0684393.
 XX
 XX (PEKE) PE CORP NY.
 XX
 XX Wei M, Zhong W, Ketchum KA, Difrancesco V, Beasley EM;
 DR WPI: 2002-444173/47.
 DR P-PSDB: AAO21501.
 XX
 XX Novel G protein coupled receptor, useful for raising antibodies, to
 PT elicit immune response, and as a reagent in assays designed to
 XX quantitatively determine levels of protein in biological samples -
 XX
 XX Claim 1; Fig 1; 82pp; English.
 XX
 XX The invention relates to an isolated G-protein coupled receptor (GPCR)
 CC polypeptide, comprising an 852 residue amino acid sequence, given in the
 CC specification, an allelic variant or ortholog of the protein, or a
 CC fragment comprising at least 10 contiguous amino acids of the protein.
 CC GPCR is useful for identifying a modulator of GPCR and an agent that
 CC binds to GPCR. GPCR and a gene chip comprising GPCR are useful as models
 CC for the development of human therapeutic agents. GPCR is useful for
 CC raising antibodies, to elicit immune response, as a reagent in assays
 CC designed to quantitatively determine levels of protein in biological
 CC samples, and as markers for tissues in which the corresponding protein
 CC is preferentially expressed. A gene chip containing GPCR is also useful
 CC as a probe for determining the chromosomal positions of nucleic acid
 CC molecules by means of in situ hybridisation. In making vectors containing
 CC the gene regulatory regions of a gene chip containing GPCR, for designing
 CC rhizomes, in making vectors that express GPCR, and for constructing host
 CC cells and transgenic animals expressing nucleic acid molecules and
 CC peptides. A host cell containing GPCR is useful for conducting cell-based
 CC assays involving GPCR protein or its fragments, and for identifying
 CC protein mutants in which these function is affected. The polynucleotide
 CC encoding GPCR can be used to treat disorders by gene therapy. This
 CC polynucleotide sequence represents cDNA encoding the human G-protein
 XX coupled receptor of the invention.
 XX
 XX Sequence 3458 BP; 590 A; 1179 C; 1092 G; 597 T; 0 other;
 S0
 Query Match 99.9%; Score 2557.4; DB 24; Length 3458;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2558; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 ATGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 60
 DB 30 ATGTGTGACACTGAGAGCGGGAGCTACCGGAAAAACCGAGACGATCGCTGACCTTT 89
 QY 61 GCCCATTTGTGCTGACAGCAACTTGTAGATGAAGGGGAGCTACGCTGTGGGGGGCTG 120
 DB 90 GCCCATTTGTGCTGACAGCAACTTGTAGATGAAGGGGAGCTACGCTGTGGGGGGCTG 149

QY 121 TTCCCTGGGCGAGGCGAGAGAGGCTGGCTCCGCAACCCGAGACGCGCCAGAGCCCT 180
Db 150 TTCCCTGGGCGAGGCGAGAGAGGCTGGCTCCGCAACCCGAGACGCGCCAGAGCCCT 209
QY 181 GTGGAGACAGAGTTCTCTCAAAAGGCGCTGCTGGGAGCTGGGAGCTGAAATAATGCGCGTG 240
Db 210 GTGGAGACAGAGTTCTCTCAAAAGGCGCTGCTGGGAGCTGGGAGCTGAAATAATGCGCGTG 269
QY 241 GAGAGATCAACAAGATGCGATGCTGGCGGAGCTGGCGGAGCTGGGAGCTGGGAGCTGGG 300
Db 270 GAGAGATCAACAAGATGCGATGCTGGCGGAGCTGGCGGAGCTGGGAGCTGGGAGCTGGG 329
QY 301 GATACGCTGCTGGAGGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCA 360
Db 330 GATACGCTGCTGGAGGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCA 389
QY 361 GGCAGCGCGGACATGCGCGCTTACTGCAATACACGAGTACACGAGTACACGAGTACACGAGTAC 420
Db 390 GGCAGCGCGGACATGCGCGCTTACTGCAATACACGAGTACACGAGTACACGAGTACACGAGTAC 449
QY 421 GTCAATCGGCGGCGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 480
Db 450 GTCAATCGGCGGCGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 509
QY 481 CTCAATGCGGCGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 540
Db 510 CTCAATGCGGCGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 569
QY 541 CCCCTCTCTTCCGCGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 600
Db 570 CCCCTCTCTTCCGCGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 629
QY 601 CTGAGAGAGTTCGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 660
Db 630 CTGAGAGAGTTCGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 689
QY 661 CAGGAGCTGAGAGCTTCCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 720
Db 690 CAGGAGCTGAGAGCTTCCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 749
QY 721 GGCCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 780
Db 750 GGCCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 809
QY 781 CACCAAGTGAACAGAGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 840
Db 810 CACCAAGTGAACAGAGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 869
QY 841 CACGAGCTTCACACTACAGATACAGAGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGG 900
Db 870 CACGAGCTTCACACTACAGATACAGAGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGG 929
QY 901 GAGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 960
Db 930 GAGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 989
QY 961 GTGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 1020
Db 990 GTGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 1049
QY 1021 CACCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 1080
Db 1050 CACCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 1109
QY 1081 CTGAGAGAGAGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 1140
Db 1110 CTGAGAGAGAGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 1169
QY 1141 GTGAGAGAGAGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 1200
Db 1170 GTGAGAGAGAGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCT 1229

QY 1201 GCCAGAGGCGCTGACAACTCTTACAGTACAGAGCTTCAGGCTGCGCGGAGAGCC 1260
Db 1230 GCCAGAGGCGCTGACAACTCTTACAGTACAGAGCTTCAGGCTGCGCGGAGAGCC 1289
QY 1261 GTGAAGGCGCTGACAGCTTCAGAGAGCTTCAGAGAGCTTCAGAGAGCTTCAGAGAGCT 1320
Db 1290 GTGAAGGCGCTGACAGCTTCAGAGAGCTTCAGAGAGCTTCAGAGAGCTTCAGAGAGCT 1349
QY 1321 CCGCTGGGAGTTCAGACAG 1380
Db 1350 CCGCTGGGAGTTCAGACAG 1409
QY 1381 TGCGAGGAGCTCAGTGGCCAGGCTCCAGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCA 1440
Db 1410 TGCGAGGAGCTCAGTGGCCAGGCTCCAGAGCTGGGAGCTGGGAGCTGGGAGCTGGGAGCA 1469
QY 1441 GAGGAGCTGAAAGATCCGCTGGCAGAGCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1500
Db 1470 GAGGAGCTGAAAGATCCGCTGGCAGAGCTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1529
QY 1501 CGGAGAGTGGCAG 1560
Db 1530 CGGAGAGTGGCAG 1589
QY 1561 TGTGTGAGCTGAG 1620
Db 1590 TGTGTGAGCTGAG 1649
QY 1621 TGTGTGAG 1680
Db 1650 TGTGTGAG 1709
QY 1681 TTCTGGAGATGGGAG 1740
Db 1710 TTCTGGAGATGGGAG 1769
QY 1741 GGCCTTGTGCTGGGAG 1800
Db 1770 GGCCTTGTGCTGGGAG 1829
QY 1801 GCTCTGGGAG 1860
Db 1830 GCTCTGGGAG 1889
QY 1861 GTCTCTGTCTCTCTGGGAG 1920
Db 1890 GTCTCTGTCTCTCTGGGAG 1949
QY 1921 CACCTGCGGCTCAG 1980
Db 1950 CACCTGCGGCTCAG 2009
QY 1981 GAGTCAAG 2040
Db 2010 GAGTCAAG 2069
QY 2041 GCTTGGCTGTGGTGTCTGGGAG 2100
Db 2070 GCTTGGCTGTGGTGTCTGGGAG 2129
QY 2101 GTGGAGCTTCCCGGAG 2160
Db 2130 GTGGAGCTTCCCGGAG 2189
QY 2161 CACTGCGGAG 2220
Db 2190 CACTGCGGAG 2249
QY 2221 GCTTCTTCTGTCTCTCTGGGAG 2280
Db 2250 GCTTCTTCTGTCTCTCTGGGAG 2309
QY 2281 GCCCGAGCTCAGCTTGGCAG 2340

QY	901	GAGGCGTGGTGACCTCTGACCTGCGTCATGCGGGCTGCCGGATGAGCCAGTGGGACG	960
Db	901	GAGGCGTGGTGACCTCTGACCTGCGTCATGCGGGCTGCCGGATGAGCCAGTGGGACG	960
QY	961	GTCCTTGCGCTTCTCCAGAGGGGTGCCAGCTGCACAGAGTTCCTCCAGTACGTGAAGC	1020
Db	961	GTCCTTGCGCTTCTCCAGAGGGGTGCCAGCTGCACAGAGTTCCTCCAGTACGTGAAGC	1020
QY	1021	CACCTGGCCCTGGGCAACCGACCCGGCCCTTGCTGACCTGGGGAGAGGGAGAGAGGT	1080
Db	1021	CACCTGGCCCTGGGCAACCGACCCGGCCCTTGCTGACCTGGGGAGAGGGAGAGAGGT	1080
QY	1081	CTGAGAGAGAGACGTGTGGGCCAGCGCTGCCCGCAGTGTGACTGCATCAGCTGCAGAAC	1140
Db	1081	CTGAGAGAGAGACGTGTGGGCCAGCGCTGCCCGCAGTGTGACTGCATCAGCTGCAGAAC	1140
QY	1141	GTCGAGCCGACGGCTAAATATCAACACAGACGTTCTCTGTACGCAAGCTGTATAGCTG	1200
Db	1141	GTCGAGCCGACGGCTAAATATCAACACAGACGTTCTCTGTGTAGCGAGCTGTATAGCTG	1200
QY	1201	GCCGAGGCGCCTGACAACTCTTTCAGTCAACGCGCTCAGCGTCCCGCGCAGAGACCC	1260
Db	1201	GCCGAGGCGCCTGACAACTCTTTCAGTGTCAACGCTCAGCGTCCCGCGCAGAGACCC	1260
QY	1261	GTCGAGCCCTTGCGACGCTCTGTGAGAACATGTACAACTGACCTTCACAGTGGGGGGCTG	1320
Db	1261	GTCGAGCCCTTGCGACGCTCTGTGAGAACATGTACAACTGACCTTCACAGTGGGGGGCTG	1320
QY	1321	CCGCTGGGGTTCGACAGCACCGGAAACGTGGACATGAGATACGCTGAGACGCTGGGGTG	1380
Db	1321	CCGCTGGGGTTCGACAGCACCGGAAACGTGGACATGAGATACGACCTGAGACGCTGGGGTG	1380
QY	1381	TGGCAGGGCTCAETGGCCAGGCTCCACGACGTGGGAGGTTTCAMCGCAGGCTCAGACA	1440
Db	1381	TGGCAGGGCTCAETGGCCAGGCTCCACGACGTGGGAGGTTTCAMCGCAGGCTCAGACA	1440
QY	1441	GAGGCGCTGAAAGTCCGCTGGGCAACGCTGTGACAAACCAACCCCGTGTCCCGGGCTCG	1500
Db	1441	GAGGCGCTGAAAGTCCGCTGTGGCAACGCTGTGACAAACCAACCCCGTGTCCCGGGCTCG	1500
QY	1501	CGGCAAGTGCAGAGAGGGCCAGGTGCCCGGAGCAAGACGCGGATTCACCTCGCTCAGAC	1560
Db	1501	CGGCAAGTGCAGAGAGGGCCAGGTGGCCCGGAGCAAGACGCGGATTCACCTCGCTCAGAC	1560
QY	1561	TGTGTGACCTGCGAGGCGGGCAGCTCAGCGGAAACCCAGACGACATCGCTGCACCTTT	1620
Db	1561	TGTGTGACCTGCGAGGCGGGCAGCTCAGCGGAAACCCAGAGACATCGCTGCACCTTT	1620
QY	1621	TGTGGCCAGATGATGTGTCTCCCGGAGCAAGACACGCTGTTCCGCCGCAAGGTTCTGG	1680
Db	1621	TGTGGCCAGATGATGTGTCTCCCGGAGCAAGACACGCTGTTCCGCCGCAAGGTTCTGG	1680
QY	1681	TTCCTGCGATGAGGAGCGAGCCGCGTGTGTGTGTCTGCTCTGAGCCTGGCGCTG	1740
Db	1681	TTCCTGCGATGAGGAGCGAGCGCGTGTGTGTGTGTCTGCTCTGAGCCTGGCGCTG	1740
QY	1741	GCGCTTGTGCTGTCTGTTGGGGGCTGTCTGTACACATCGGGAAGAGCCCACTGGTTAG	1800
Db	1741	GCGCTTGTGCTGTCTGTTGGGGGCTGTCTGTACACATCGGGAAGAGCCCACTGGTTAG	1800
QY	1801	GCGCTGGGGGGGCGCCTGCGGCTGTTTGGACTGTGTGTGTGTGGGCTGTGTGTGCTCAGC	1860
Db	1801	GCGCTGGGGGGGCGCCTGCGGCTGTCTTGGGCTGTGTGTGTGTGGGCTGTGTGTGCTCAGC	1860
QY	1861	GTCCTCTGTTCCCTTGCGCAGCCAGCCCTTGCCATGTGCTGGCCACAGACCTTGTTC	1920
Db	1861	GTCCTCTGTTCCCTTGCGCAGCCAGCCCTTGCCATGTGCTGGCCACAGACCTTGTTC	1920
QY	1921	CACCTCCCGCTCACGGGCTCCGTGAACACACTTCTCTGACAGGGCGGCGAATCTTGCTG	1980
Db	1921	CACCTCCCGCTCACGGGCTCCGTGAACACACTTCTCTGACAGGGCGGCGAATCTTGCTG	1980

QY	1981	GAGCACAATGCTGCTGACCTGGGACAGCCGGCTAGAGCTGCTCGGGGGCCCTGG	2040
Db	1981	GAGTCAGAACTGCTCTGAGCTGGGACAGCCGGCTAGAGCTGCTCGGGGGCCCTGG	2040
QY	2041	GCCGTGCTGGTGGTGTCTGGCCATGCTGTGGAGGTCGCACCTGTGACCTGTGACTCG	2100
Db	2041	GCCGTGCTGGTGGTGTCTGGCCATGCTGTGGAGGTCGCACCTGTGACCTGTGACTCG	2100
QY	2101	GTGGGCTTCCCGCGCGAGGTGGTGGAGAGCTGGACATGCTGTCCCGAGAGGCGTGGTG	2160
Db	2101	GTGGGCTTCCCGCGCGAGGTGGTGGAGAGCTGGACATGCTGTCCCGAGAGGCGTGGTG	2160
QY	2161	CACATCCGACACAGCTCTCTGGGTCACTTCGGGCTTAGCCAGCCAGCAATGCCACGCTG	2220
Db	2161	CACATCCGACACAGCTCTCTGGGTCACTTCGGGCTTAGCCAGCCAGCAATGCCACGCTG	2220
QY	2221	GCCCTTCTCTGCTTCTCTGGGCACTTTCCTGTGGGAGCCAGCCGGGCTGCTACAAACGT	2280
Db	2221	GCCCTTCTCTGCTTCTCTGGGCACTTTCCTGTGGGAGCCAGCCGGGCTGCTACAAACGT	2280
QY	2281	GCCCGGTGGCCTCACTTTTGGCATGCTGGCTTACTATACCTGGGTCTCCTTTGTGCC	2340
Db	2281	GCCCGGTGGCCTCACTTTTGGCATGCTGGCTTACTATACCTGGGTCTCCTTTGTGCC	2340
QY	2341	CTCCTGGCCAAATGAGAGGTGTCCTCAGAGCCGCGCTGCAAAATGGGGCCCTCTGCTC	2400
Db	2341	CTCCTGGCCAAATGAGAGGTGTCCTCAGAGCCGCGCTGCAAAATGGGGCCCTCTGCTC	2400
QY	2401	TGTGTCTCTGGGACCTCTGGCTGCGCTTCACACTGCCAGGTGTATTACCTGCTCATGTGGCAG	2460
Db	2401	TGTGTCTCTGGGACCTCTGGCTGCGCTTCACACTGCCAGGTGTATTACCTGCTCATGTGGCAG	2460
QY	2461	CCAGGGCTCAACACCCCCAGATTCTTCTGGGAGGGGGCCCTGGGGATATGCCAAGGCCAG	2520
Db	2461	CCAGGGCTCAACACCCCCAGATTCTTCTGGGAGGGGGCCCTGGGGATATGCCAAGGCCAG	2520
QY	2521	AATGAGGGGAACAGAGAAATCAGGGGGAACCAATGATGA	2559
Db	2521	AATGAGGGGAACAGAGAAATCAGGGGGAACCAATGATGA	2559

RESULT 4
AAS14574
ID AAS14574 standard; cDNA; 3489 BP

AC	AAS14574;
XX	
DT	18-DEC-2001 (first entry)
DT	

DE Human cDNA encoding G protein-coupled receptor, GPCR, 50289.

KM Human;ass: GPCR; g protein-coupled receptor; 50289; cardiant;
KM antiatherosclerotic; analgesic; cytoskeletal; antianginal;
KM cardiovascular disorder; angiogenesis-related disorder;
KM neural disorder: pain response disorder; inflammatory disorder;
KM atherosclerosis; angina pectoris; myocardial infarction;
KM ischaemic h heart disease; sudden cardiac death; obesity;
KM ischaemic heart disease; diabetes; prostate cancer-related pain

OS Homo sapiens

FH	Key	Location/Qualifiers
1	1	1
2	2	2
3	3	3
4	4	4
5	5	5
6	6	6
7	7	7
8	8	8
9	9	9
10	10	10
11	11	11
12	12	12
13	13	13
14	14	14
15	15	15
16	16	16
17	17	17
18	18	18
19	19	19
20	20	20
21	21	21
22	22	22
23	23	23
24	24	24
25	25	25
26	26	26
27	27	27
28	28	28
29	29	29
30	30	30
31	31	31
32	32	32
33	33	33
34	34	34
35	35	35
36	36	36
37	37	37
38	38	38
39	39	39
40	40	40
41	41	41
42	42	42
43	43	43
44	44	44
45	45	45
46	46	46
47	47	47
48	48	48
49	49	49
50	50	50
51	51	51
52	52	52
53	53	53
54	54	54
55	55	55
56	56	56
57	57	57
58	58	58
59	59	59
60	60	60
61	61	61
62	62	62
63	63	63
64	64	64
65	65	65
66	66	66
67	67	67
68	68	68
69	69	69
70	70	70
71	71	71
72	72	72
73	73	73
74	74	74
75	75	75
76	76	76
77	77	77
78	78	78
79	79	79
80	80	80
81	81	81
82	82	82
83	83	83
84	84	84
85	85	85
86	86	86
87	87	87
88	88	88
89	89	89
90	90	90
91	91	91
92	92	92
93	93	93
94	94	94
95	95	95
96	96	96
97	97	97
98	98	98
99	99	99
100	100	100

```
/*tag= a
```

/note= "This region is specifically claimed in

FT	siq_peptide
1	1
2	1
3	1
4	1
5	1
6	1
7	1
8	1
9	1
10	1
11	1
12	1
13	1
14	1
15	1
16	1
17	1
18	1
19	1
20	1
21	1
22	1
23	1
24	1
25	1
26	1
27	1
28	1
29	1
30	1
31	1
32	1
33	1
34	1
35	1
36	1
37	1
38	1
39	1
40	1
41	1
42	1
43	1
44	1
45	1
46	1
47	1
48	1
49	1
50	1
51	1
52	1
53	1
54	1
55	1
56	1
57	1
58	1
59	1
60	1
61	1
62	1
63	1
64	1
65	1
66	1
67	1
68	1
69	1
70	1
71	1
72	1
73	1
74	1
75	1
76	1
77	1
78	1
79	1
80	1
81	1
82	1
83	1
84	1
85	1
86	1
87	1
88	1
89	1
90	1
91	1
92	1
93	1
94	1
95	1
96	1
97	1
98	1
99	1
100	1

mat. peptide

```

F1      / *tag= c
Fm      /]abel= Mature GPCR50289

```

XX

PN	MO200164882-A2.
XX	
PD	07-SEP-2001.
XX	
PF	28-FEB-2001; 2001WO-US06543.
XX	
PR	29-FEB-2000; 2000US-186059P.
XX	
PA	(MILL-) MILLENNIUM PHARM INC.
XX	
PI	Glucksmann MA, Galvin KM, Silos-Santlago I;
XX	
DR	WPI: 2001-589866/66.
DR	P-PDSB: AAU08396.
XX	
PT	Novel G protein coupled receptors and nucleic acids encoding them, for identifying agents for the treatment of cardiac disorders -
PS	Claim 2; Fig 15; 209pp; English.
XX	
CC	The invention relates to novel human G protein-coupled receptors (GPCR)
CC	named 1983, 52881, 2398, 45449, 50289 and 52872. The novel GPCRs and
CC	nucleic acids encoding them are useful for identifying agents for the
CC	treatment of cardiovascular disorders, angioneurosis-related disorders,
CC	neural disorders, pain response disorders and inflammatory disorders,
CC	e.g. atherosclerosis, angina pectoris and myocardial infarction,
CC	ischemic heart disease, sudden cardiac death, hypertensive heart
CC	disease, diabetes, prostate cancer-related pain, diabetes and obesity.
CC	The present sequence encodes GPCR 50289.
XX	
SQ	Sequence 3489 BP; 578 A; 1189 C; 1102 G; 612 T; 8 other;
	Query Match 99.9%; Score 2555.8; DB 22; Length 3489; Best Local Similarity 99.9%; Pred. No. 0; Matches 2557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY	1 ATGCTGGGGCCCTGCCTCTCTGGGCGCTCAAGCCTTGTGGCTCTCTGACCCCTGGACGGG 60
DB	52 ATGCTGGGGCCCTGCCTCTCTGGGCGCTCAAGCCTTGTGGCTCTCTGACCCCTGGACGGG 111
OY	61 GCCCATTTGTGCTCTCACAGCAACTTAGATGAAGGGGACTACGTGCGGGGGCTG 120
DB	112 GCCCATTTGTGCTCTCACAGCAACTTAGATGAAGGGGACTACGTGCGGGGGCTG 171
OY	121 TTCGCCCTGGGCGAGGCGAGAGGCTGGCCCTCCGACGCCGACAGCGCCAGACGCT 180
DB	172 TTCGCCCTGGGCGAGGCGAGAGGCTGGCCCTCCGACGCCGACAGCGCCAGACGCT 231
OY	181 GTGTGCACCAAGTTTCCTCTCAAAAGCGCTGCTTGCGGCACTGCGCATGAATAATGGCGTG 240
DB	232 GTGTGCACCAAGTTTCCTCTCAAAAGCGCTGCTTGCGGCACTGCGCATGAATAATGGCGTG 291
OY	241 GAGGAGATCAACAAGAAGCGAGTATGCGTGGCCGGGGCTGGCGCTGAGACCTCTT 300
DB	292 GAGGAGATCAACAAGAAGCGAGTATGCGTGGCCGGGGCTGGCGCTGAGACCTCTT 351
OY	301 GATACGTGCTCGAGCCTGTGTGGCCAGTAAGAACCCAGCCTCATGTTCTTGCCAAAGCA 360
DB	352 GATACGTGCTCGAGCCTGTGTGGCCAGTAAGAACCCAGCCTCATGTTCTTGCCAAAGCA 411
OY	361 GCGACCGCGGACATGCGCGCTTACAGCACTACAGCAGTACACGCCCCGCTGTGCTGCT 420
DB	412 GCGACCGCGGACATGCGCGCTTACAGCACTACAGCAGTACACGCCCCGCTGTGCTGCT 471
OY	421 GTCATCGGGGCCCATCGTCAGAGTCCGCAAGTCCACCGGGAAGTTCATAGCTTCTTC 480
DB	472 GTCATCGGGGCCCATCGTCAGAGTCCGCAAGTCCACCGGGAAGTTCATAGCTTCTTC 531
OY	481 CTCATGCCCAAGTAGTACGCTCTACATGAGAGCTCTGAGCGCCCGGAGAGACCTTC 540
DB	532 CTCATGCCCAAGTAGTACGCTCTACATGAGAGCTCTGAGCGCCCGGAGAGACCTTC 591
OY	541 CCCTCCCTTTCGCGACCGTGCACGCGAAGCGGTGACGTGACGTGACGCGCCGCGGAGCTG 600

[illegible]

Db 181 GTGTGACACAGGTCTCTCAACAGGCCCTCTGTGGGCACTGTGCCATGTAAATAGGCCGTG 240
QY 241 GAGGAGATCAACAACAAGTCGGATGCTGCGCCGGGCTGGGCTGACAGCCTTT 300
Db 241 GAGGAGATCAACAACAAGTCGGATGCTGCGCCGGGCTGGGCTGACAGCCTTT 300
QY 301 GATAGGTCTCGAGGCTGTGTGGCCATGAAGCCAGCCTCATGTTCTTGGCCAGGCA 360
Db 301 GATAGGTCTCGAGGCTGTGTGGCCATGAAGCCAGCCTCATGTTCTTGGCCAGGCA 360
QY 361 GGCAGCCGGGACATGCGCCCTACTGTCAACTACAGCAGTACAGCCCGCTGTGCTG 420
Db 361 GGCAGCCGGGACATGCGCCCTACTGTCAACTACAGCAGTACAGCCCGCTGTGCTG 420
QY 421 GTCATGCGGCCCACTGCTCAGAGCTGCGCATGTGACCGGGAAGTTCTTCAGTCTTTC 480
Db 421 GTCATGCGGCCCACTGCTCAGAGCTGCGCATGTGACCGGGAAGTTCTTCAGTCTTTC 480
QY 481 GTCATGCGGCCCACTGCTCAGAGCTGCTAGCATGTGAGACCTGTGAGCCCGGGAGACCTTC 540
Db 481 GTCATGCGGCCCACTGCTCAGAGCTGCTAGCATGTGAGACCTGTGAGCCCGGGAGACCTTC 540
QY 541 CCTCTCTTCTTCCGACCGCTGCGCCAGCCGCTGTGTACACTGACGCGCCCGCGAGCTG 600
Db 541 CCTCTCTTCTTCCGACCGCTGCGCCAGCCGCTGTGTACACTGACGCGCCCGCGAGCTG 600
QY 594 CCTCTCTTCTTCCGACCGCTGCGCCAGCCGCTGTGTACACTGACGCGCCCGCGAGCTG 594
Db 594 CCTCTCTTCTTCCGACCGCTGCGCCAGCCGCTGTGTACACTGACGCGCCCGCGAGCTG 594
QY 601 CTGCAGAGATTGGGTGTAAGTGGGTGGCGCCCTGGGAGGAGACAGATACGGCCG 660
Db 595 CTGCAGAGATTGGGTGTAAGTGGGTGGCGCCCTGGGAGGAGACAGATACGGCCG 654
QY 661 CAGGGCCTGAGCATCTTCTGCGCCCTGCGCGCGGACGCGGCAATGTCATCGCGCAGAG 720
Db 655 CAGGGCCTGAGCATCTTCTGCGCCCTGCGCGCGGACGCGGCAATGTCATCGCGCAGAG 714
QY 721 GGCGTGTGGCGCTGCGCCGCTGCGCATGTGCTGGGCTGGGGAAGGTGAGAGACGCTCTG 780
Db 715 GGCGTGTGGCGCTGCGCCGCTGCGCATGTGCTGGGCTGGGGAAGGTGAGAGACGCTCTG 774
QY 781 CACAGGTGAACACAGAGCGTGTGAGGTGTGCTGTTGGCTCCCTGACGCGCGC 840
Db 775 CACAGGTGAACACAGAGCGTGTGAGGTGTGCTGTTGGCTCCCTGACGCGCGC 834
QY 841 CAGCGCCTTCTTAACATACAGCATACAGACAGAGCTCTGCGCCAAAGGTGTGGGTGGCAGC 900
Db 835 CAGCGCCTTCTTAACATACAGCATACAGACAGAGCTCTGCGCCAAAGGTGTGGGTGGCAGC 894
QY 901 GAGGCGCTGCTGACCTCTTACCTGTGTGATGGGGCTGCCCGCATGGCCCAATGGGCAAG 960
Db 895 GAGGCGCTGCTGACCTCTTACCTGTGTGATGGGGCTGCCCGCATGGCCCAATGGGCAAG 954
QY 961 GTGCTTGGCTTCTCTCAGAGGGGTGCCAGCTGACAGATTTCCCGCATGTAAAGTGAAGT 1020
Db 955 GTGCTTGGCTTCTCTCAGAGGGGTGCCAGCTGACAGATTTCCCGCATGTAAAGTGAAGT 1014
QY 1021 CACTGTGGCCCTGGCAACCGACCCGGCTTCTGCTGTGCGCTGGGAGAGAGGAGCAGGCT 1080
Db 1015 CACTGTGGCCCTGGCAACCGACCCGGCTTCTGCTGTGCGCTGGGAGAGAGGAGCAGGCT 1074
QY 1081 CTGGAGAGAGAGAGTGGTGGGCGCAGCGCTGCCGAGTGTGATGCACTGCGCAAGC 1140
Db 1075 CTGGAGAGAGAGAGTGGTGGGCGCAGCGCTGCCGAGTGTGATGCACTGCGCAAGC 1134
QY 1141 GTGAGGCGAGGGCTAAATACCAACAGAGATCTCTGTGTAGGAGCTGTGTATAGCGTG 1200
Db 1135 GTGAGGCGAGGGCTAAATACCAACAGAGATCTCTGTGTAGGAGCTGTGTATAGCGTG 1194
QY 1201 GCCCAGGGCCCTGCACAACACTCTTCACTGCAACGCTTCAAGCTGCGCCCGCAGAGACCC 1260
Db 1195 GCCCAGGGCCCTGCACAACACTCTTCACTGCAACGCTTCAAGCTGCGCCCGCAGAGACCC 1254
QY 1261 GTGAAGCCTTGGCAGCTCTCTGAGAAACATGTACAACCTGACCTTCCAGCTGGCGGGCTG 1320
Db 1255 GTGAAGCCTTGGCAGCTCTCTGAGAAACATGTACAACCTGACCTTCCAGCTGGCGGGCTG 1314

QY 1321 CCGCTGCGGTTTCAGACAGCAGCGAAACGTGACATGAGTACGACTGAAGCTGTGGTG 1380
Db 1315 CCGCTGCGGTTTCAGACAGCAGCGAAACGTGACATGAGTACGACTGAAGCTGTGGTG 1374
QY 1381 TGGCAGGGCTCAGTGGCCAGGCTCCAGCAGTGTGGCAGGTTCAACGGCAGCCTCAGGACA 1440
Db 1375 TGGCAGGGCTCAGTGGCCAGGCTCCAGCAGTGTGGCAGGTTCAACGGCAGCCTCAGGACA 1434
QY 1441 GAGCGCTTAAGATCCGCTTGGCAGACGCTGTGACAAACCAAGAGCCGTGTCCGGTCTG 1500
Db 1435 GAGCGCTTAAGATCCGCTTGGCAGACGCTGTGACAAACCAAGAGCCGTGTCCGGTCTG 1494
QY 1501 CCGCAGTGGCAGAGGGGAGGAGTGGCGCCGGGTCAGAGGAGGTTCCACTGCTGTGCTGAC 1560
Db 1495 CCGCAGTGGCAGAGGGGAGGAGTGGCGCCGGGTCAGAGGAGGTTCCACTGCTGTGCTGAC 1554
QY 1561 TGTGTGACTGCGAGGCGGGCAGCTACCGGCAAAACCCAGAGACATGCGCTGCACCTTT 1620
Db 1555 TGTGTGACTGCGAGGCGGGCAGCTACCGGCAAAACCCAGAGACATGCGCTGCACCTTT 1614
QY 1621 TGTGGCAGAGTATGATGTGTCCTCCGAGCGCAAGCAGCTGCTTCCGCGGAGTCTCG 1680
Db 1615 TGTGGCAGAGTATGATGTGTCCTCCGAGCGCAAGCAGCTGCTTCCGCGGAGTCTCG 1674
QY 1681 TTCCTGCAATGGGCGAGCGGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1740
Db 1675 TTCCTGCAATGGGCGAGCGGCTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1734
QY 1741 GGCCTTGTGCTGCTGCTTGTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
Db 1735 GGCCTTGTGCTGCTGCTTGTGGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1794
QY 1801 GCGTGGGGGGGCGCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
Db 1795 GCGTGGGGGGGCGCCCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1854
QY 1861 GTCTTCCTGCTTCCCTGCGGAGCCAGCCCTGCGCGAGCTGCTGCGCCAGAGCCTTGTGC 1920
Db 1855 GTCTTCCTGCTTCCCTGCGGAGCCAGCCCTGCGCGAGCTGCTGCGCCAGAGCCTTGTGC 1914
QY 1921 CACTCCCGCTCAAGGGCTGCTGAGCAGACACTTCTCTGAGGCGGCGCAGATCTTCTG 1980
Db 1915 CACTCCCGCTCAAGGGCTGCTGAGCAGACACTTCTCTGAGGCGGCGCAGATCTTCTG 1974
QY 1981 GAGTCAAGTCTGCTTGTGAGCTGTGGGAGAGACCGGCTGAGTGGCTGCTGCGGGGCGCTG 2040
Db 1975 GAGTCAAGTCTGCTTGTGAGCTGTGGGAGAGACCGGCTGAGTGGCTGCTGCGGGGCGCTG 2034
QY 2041 GCGTGGCTGTGGTGTGCTGCGCATGTGAGAGTGGCACTGTGCACTGTGCTGCTG 2100
Db 2035 GCGTGGCTGTGGTGTGCTGCGCATGTGAGAGTGGCACTGTGCACTGTGCTGCTG 2094
QY 2101 GTGGCCTTCCCGGAGGTGTGAGCAGATGCGACATGCTGTGCCAAGAGGCGCTG 2160
Db 2095 GTGGCCTTCCCGGAGGTGTGAGCAGATGCGACATGCTGTGCCAAGAGGCGCTG 2154
QY 2161 CACTGCGGACAGCCTCTGCGGTGAGCTTGGCTGTGCGCAGCGCCAAATGCGACGCTG 2220
Db 2155 CACTGCGGACAGCCTCTGCGGTGAGCTTGGCTGTGCGCAGCGCCAAATGCGACGCTG 2214
QY 2221 GCGTTTCTGTGCTTCTGCGCACTTTCCTGTGCGAGAGCAGCGGCGCTGTCAACCGT 2280
Db 2215 GCGTTTCTGTGCTTCTGCGCACTTTCCTGTGCGAGAGCAGCGGCGCTGTCAACCGT 2274
QY 2281 GCGCGTGGCTCACTTTCCTGATGCTGAGCTTTCATCATCAGCTGGGTCTCTTGTGCGC 2340
Db 2275 GCGCGTGGCTCACTTTCCTGATGCTGAGCTTTCATCATCAGCTGGGTCTCTTGTGCGC 2334
QY 2341 CTCCTGGCCAATGTGAGTGTGCTCTCAGGCGCGCGCTGCAATGTGGCGCCTCTGCTG 2400
Db 2335 CTCCTGGCCAATGTGAGTGTGCTCTCAGGCGCGCGCTGCAATGTGGCGCCTCTGCTG 2394

QY 2401 TGTGTCTGGGATCTGGCTGCTTCACCTGCCAGGTGTACTCTCATGCGGAG 2460
 Db 2395 TGTGTCTGGGATCTGGCTGCTTCACCTGCCAGGTGTACTCTCATGCGGAG 2454
 QY 2461 CCAGGGCTCAACACCCCGAGTTCTCTCTGGAGGGGCGCTGGGATGCCAAGGCCAG 2520
 Db 2455 CCAGGGCTCAACACCCCGAGTTCTCTCTGGAGGGGCGCTGGGATGCCAAGGCCAG 2514
 QY 2521 AATGACGGGACACAGAAATCAGGGGAAACATGAGTGA 2559
 Db 2515 AATGACGGGACACAGAAATCAGGGGAAACATGAGTGA 2553

RESULT 6
 ABL55952
 ID ABL55952 standard; cDNA; 2729 BP.
 AC ABL55952;
 XX
 DT 17-JUN-2002 (first entry)
 XX
 DE Human G-protein coupled receptor encoding cDNA SEQ ID NO 12.
 XX
 KW Human; GPCR; G-protein coupled receptor; receptor; anti-HIV; antitumour;
 KW antiinflammatory; antiallergic; antianemic; antiasthmatic; antiviral;
 KW immunosuppressive; dermatological; nephrotropic; antigout; antihypertoid;
 KW cytosolic; neuroprotective; osteopathic; antipsoriatic; antirheumatic;
 KW antiallergic; thymometric; antitumor; ophthalmological; antibacterial;
 KW fungicide; antiparasitic; protozoacide; antihelminthic; antidiabetic;
 KW antileptocytic; hepatotropic; anticonvulsant; anorectic; metabolic;
 KW antiemetic; antidiarrhoeic; neuroleptic; cerebroprotective; nocotropic;
 KW antiparkinsonian; depilatory; tranquilizer; hypotensive; vasotropic;
 KW cardiant; antidiagonal; vulnerary; proliferative disorder; cancer;
 KW neurological disorder; Alzheimer's disease; Huntington's disease;
 KW Parkinson's disease; multiple sclerosis; meningitis; prion;
 KW cardiovascular disorder; acquired immunodeficiency syndrome; AIDS;
 KW Crohn's disease; diabetes mellitus; rheumatoid arthritis; gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..2592
 FT /*tag= a
 FT /product= "G-protein coupled receptor"
 FT
 PN MO200198323-A2.
 XX
 PD 27-DEC-2001.
 XX
 PE 15-JUN-2001; 2001MO-US19354.
 XX
 PR 16-JUN-2000; 2000US-212483P.
 PR 23-JUN-2000; 2000US-213950P.
 PR 26-JUN-2000; 2000US-214062P.
 PR 07-JUL-2000; 2000US-216595P.
 PR 14-JUL-2000; 2000US-218936P.
 PR 19-JUL-2000; 2000US-219154P.
 XX
 PA (INCYTE GENOMICS INC.
 XX
 PI Lal P, Gnanu R, Hafalia AJA, Walla NK, Thornton M, Nguyen DB;
 PI Lu Y, Gandhi AR, Patterson C, Kallilik DA, Baughn MR, Ramkumar J;
 PI Triboley CM, Lee EA, Ding L, Burford N, Yao MG, Yang J;
 PI Griffin JA;
 XX
 DR WPI: 2002-139780/18.
 DR P-PSDB: ABB77318.
 XX
 PT Novel G-protein coupled receptor protein and polynucleotides useful for
 PT diagnosing, treating or preventing disorders of cell proliferation e.g.
 PT cancer, neurological and genetic disorder e.g. thalassemia -
 XX
 PS Claim 48; Page 116-117; 121pp; English.

XX The invention relates to a G-protein coupled receptor protein (GPCR)
 CC polypeptide (ABB77317-ABB77326). The GPCR is useful for screening an
 CC agonist/antagonist of GPCR, a compound that specifically binds to GPCR
 CC or that modulates the activity of GPCR. GPCR is also useful as an
 CC immunogen for preparing antibodies which are useful for diagnosing a
 CC condition of disease associated with expression of GPCR in a subject,
 CC for detecting and purifying GPCR from a sample. The GPCR encoding
 CC polynucleotide (ABL55951-ABL55960) is useful for screening for a compound
 CC effective in altering expression of GPCR. GPCR is also useful for
 CC assessing toxicity of a test compound useful for treating a disease or
 CC conditional associated with decreased expression or overexpression of
 CC functional GPCR. Examples of disorders include cell proliferative
 CC disorder such as arteriosclerosis, atherosclerosis, hepatitis, mixed
 CC connective tissue disease (MCTD), psoriasis and cancer including
 CC adenocarcinoma, leukaemia; a neurological disorder such as epilepsy,
 CC stroke, Alzheimer's disease, Huntington's disease, dementia, Parkinson's
 CC disease, retinitis pigmentosa, multiple sclerosis, bacterial and viral
 CC meningitis, abscess, subdural empyema; prion disease including kuru,
 CC Creutzfeldt-Jakob disease, fatal familial insomnia, neurofibromatosis,
 CC tuberous sclerosis, cerebral palsy, polymyositis; inherited, metabolic,
 CC endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis;
 CC mental disorders including mood, anxiety, and schizophrenic disorders;
 CC seasonal affective disorder (SAD); akathisia, amnesia, cataplexy,
 CC diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses,
 CC Tourette's disorder; cardiovascular disorders such as hypertension,
 CC vasculitis, vascular tumours, congestive heart failure, ischaemic heart
 CC disease, myocardial infarction, calcific aortic valve stenosis,
 CC infective endocarditis, endocarditis of systemic lupus erythematosus,
 CC cardiac transplantation; gastrointestinal disorder such as dysphagia,
 CC gastritis, anorexia, nausea, emesis, abdominal angina, infections of
 CC the intestinal tract, peptic ulcer, hepatitis, cirrhosis, diarrhoea,
 CC acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, Reye's
 CC syndrome, liver infarction; an autoimmune/inflammatory disorder such as
 CC AIDS, Addison's disease, adult respiratory distress syndrome,
 CC allergies, amyloidosis, anaemia, asthma, atherosclerosis, autoimmune
 CC thyroiditis, bronchitis, Crohn's disease, diabetes mellitus,
 CC Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis,
 CC multiple sclerosis, myasthenia gravis, myocardial or pericardial
 CC inflammation, osteoporosis, psoriasis, rheumatoid arthritis,
 CC scleroderma, Sjogren's syndrome, systemic lupus erythematosus, uveitis,
 CC viral, bacterial, fungal, parasitic, protozoal, and helminthic
 CC infections, and trauma; a metabolic disorders such as diabetes,
 CC obesity, and osteoporosis. The GPCR encoding polynucleotide is useful
 CC for gene therapy and for creating knock in humanised animals (pigs) or
 CC transgenic animals (mice or rats) to model human disease.
 XX
 SO Sequence 2729 BP; 413 A; 934 C; 879 G; 503 T; 0 other;

Query Match 98.38; Score 2514.4; DB 24; Length 2729;
 Best Local Similarity 98.78; Pred. No. 0;
 Matches 2558; Conservative 0; Mismatches 1; Indels 33; Gaps 1;

QY 1 ATGTGGGGCCCTGCTGCTGCTGGGCTCAGCTCGGGCTCCGACCCCTGGAGCGGG 60
 Db 1 ATGTGGGGCCCTGCTGCTGCTGGGCTCAGCTCGGGCTCCGACCCCTGGAGCGGG 60
 QY 61 GCCCATTTGCTCTGTCAACAGCACTTGAAGTGAAGGGGAGTCACTGCTGGGGGCTG 120
 Db 61 GCCCATTTGCTCTGTCAACAGCACTTGAAGTGAAGGGGAGTCACTGCTGGGGGCTG 120
 QY 121 TTCCCTGGGGGAGGCGGAGAGGCTGCTCCGACCGGAGACGCGCCAGAGCCCT 180
 Db 121 TTCCCTGGGGGAGGCGGAGAGGCTGCTCCGACCGGAGACGCGCCAGAGCCCT 180
 QY 181 GTGTGACACAGGCTCTCTCAAAAGCGCTGCTGGGACATGCGCAATGAATGCGCGTG 240
 Db 181 GTGTGACACAGGCTCTCTCAAAAGCGCTGCTGGGACATGCGCAATGAATGCGCGTG 240
 QY 241 GAGAGATCAACACAGATGCTGCTGCGCGGGCTGGCGGCTGAGACCTCTTT 300
 Db 241 GAGAGATCAACACAGATGCTGCTGCGCGGGCTGGCGGCTGAGACCTCTTT 300

QY	301	GATACGTGCTCGAGAGCTGTGTGGTGGCCATGAAAGCCAGCCTCATGTCTCTGGCCAAAGCA	360
Dp	301	GATACGTGCTCGAGAGCTGTGTGGTGGCCATGAAAGCCAGCCTCATGTCTCTGGCCAAAGCA	360
QY	361	GGCAGCGCGCATTCGCGCGGCTACTGCACTACCTACGACAGATACACGCCCGGTGCTGGCT	420
Dp	361	GGCAGCGCGCATTCGCGCGGCTACTGCACTACCTACGACAGATACACGCCCGGTGCTGGCT	420
QY	421	GTCACTGGGGCCCCACTGCTAGAGACTGCGCATGTGTACCGGCAAGTTCTTCAAGTTCTTC	480
Dp	421	GTCACTGGGGCCCCACTGCTAGAGACTGCGCATGTGTACCGGCAAGTTCTTCAAGTTCTTC	480
QY	481	CTCATGCCCCCAGAGTCACTACGCTGTGTACATGTAGACTGTGAGCGCCCGGAGACCTTC	540
Dp	481	CTCATGCCCCCAGAGTCACTACGCTGTGTACATGTAGACTGTGAGCGCCCGGAGACCTTC	540
QY	541	CCCTCCTTCTTCGCAACCGGCCACGACCGGTGTACCTGACGAGGCCCGCGCGAGAGTG	600
Dp	541	CCCTCCTTCTTCGCAACCGGCCACGACCGGTGTACCTGACGAGGCCCGCGCGAGAGTG	600
QY	601	CTCGAGAGATTCGGCTGTGAACTGGGTGGCGGCCCTGTGGAGAGGAGACAGAGTACGGCCG	660
Dp	601	CTCGAGAGATTCGGCTGTGAACTGGGTGGCGGCCCTGTGGAGAGGAGACAGAGTACGGCCG	660
QY	661	CAGGGCCTGAGACATCTTCTGGCCCTGGCCCGGGCAGCGCGCATCTGCATCGCGCAGAG	720
Dp	661	CAGGGCCTGAGACATCTTCTGGCCCTGGCCCGGGCAGCGCGCATCTGCATCGCGCAGAG	720
QY	721	GGCCTGTGTGCGCGTCCGCCGTGCCCATGTGACTGGCGCTGTGGGAAAGTGCAGAGAGTCTTG	780
Dp	721	GGCCTGTGTGCGCGTCCGCCGTGCCCATGTGACTGGCGCTGTGGGAAAGTGCAGAGAGTCTTG	780
QY	781	CACCAAGTGAACAGAGACATGTGCAGAGGTGTGTCTTTGGCTCGTACAGCCGCG	840
Dp	781	CACCAAGTGAACAGAGACATGTGCAGAGGTGTGTCTTTGGCTCGTACAGCCGCG	840
QY	841	CAGGCCCTCTTCAACTACAGCATCAGCAGAGGCTCTCGCCCAAGGTGTGGGTGGCAGC	900
Dp	841	CAGGCCCTCTTCAACTACAGCATCAGCAGAGGCTCTCGCCCAAGGTGTGGGTGGCAGC	900
QY	901	GAGGCTGTGAGACTTGTGACTGTGTATGTGGGCTGCCCGATGGCCATGAGATGGGCAAG	960
Dp	901	GAGGCTGTGAGACTTGTGACTGTGTATGTGGGCTGCCCGATGGCCATGAGATGGGCAAG	960
QY	961	GTGCTTGGCTTCTCCAGAGGGGTGCCAGCTGCACAGATTTCCCCAGTACGTGAAGAC	1020
Dp	961	GTGCTTGGCTTCTCCAGAGGGGTGCCAGCTGCACAGATTTCCCCAGTACGTGAAGAC	1020
QY	1021	CACCTGGCCCTGTGGCACAAGCCCGGCTTGTGCTGTGCCCTGGGAGAGAGGAGAGCGGT	1080
Dp	1021	CACCTGGCCCTGTGGCACAAGCCCGGCTTGTGCTGTGCCCTGGGAGAGAGGAGAGCGGT	1080
QY	1081	CTGAGAGAGAGACGTGGTGGGCCAGGCGCTGCCGCAGTGTGACTGCATACGCTGCAGAAC	1140
Dp	1081	CTGAGAGAGAGACGTGGTGGGCCAGGCGCTGCCGCAGTGTGACTGCATACGCTGCAGAAC	1140
QY	1141	GTGAGCCAGAGGCTAAATTCACACACAGAGTTCTGTGTACGACGCTGTGTATAGCTG	1200
Dp	1141	GTGAGCCAGAGGCTAAATTCACACACAGAGTTCTGTGTACGACGCTGTGTATAGCTG	1200
QY	1201	GCCAGAGCCCTGCACAACACTCTTTCAGTGCACAAGCCTCAGGCTGCCCGCGCAGAGACCC	1260
Dp	1201	GCCAGAGCCCTGCACAACACTCTTTCAGTGCACAAGCCTCAGGCTGCCCGCGCAGAGACCC	1260
QY	1261	GTGAAGCCCTGTGGCAGCTCTCTGTGGAGAACATGTATACCTTCCAGTGTGGCGGGCTG	1320
Dp	1261	GTGAAGCCCTGTGGCAGCTCTCTGTGGAGAACATGTATACCTTCCAGTGTGGCGGGCTG	1320
QY	1321	CCGCTGGGTTTCGACACACACCGGAAACGTGGACATGTAGTACGACTTGAAGCTGTGGGTG	1380
Dp	1321	CCGCTGGGTTTCGACACACACCGGAAACGTGGACATGTAGTACGACTTGAAGCTGTGGGTG	1380
QY	1381	TGGCAGGGCTCAGTGTGCCAGGCTTCACAGACTGGGGAGGTTTCACGGGAGCCTCAGAGCA	1440

[illegible]

Db	2461	CACGTCGCCAGGCGTACCGTCTCAGCGCGCACGACGAGGCGTCAACACCCCGAGTCTTTC	2520
OY	2488	CTGGGAGGAGGGGCGCCCTGGGATGCCCCAAGGCCAATAATGACGGGAACACAGGAATCAGGGG	2547
Db	2521	CTGGGGAGGGGGCGCCCTGGGATGCCCCAAGGCCAATAATGACGGGAACACAGGAATCAGGGG	2580
OY	2548	AAACATGAGTGA	2559
Db	2581	AAACATGAGTGA	2592
RESULT 7			
ABK1615	ABK1615 standard; cDNA; 2739 BP.		
XX	ABK1615;		
XX	14-MAR-2002 (first entry)		
DE	Human G-coupled receptor (GCRC) cDNA, Seq ID No 24.		
XX	Human; cytosolic; neuroprotective; immunosuppressant; nootropic;		
KM	anti-inflammatory; anti-viral; gastrointestinal; cardiovascular;		
KM	cerebroprotective; G-coupled receptor; cell proliferative disease;		
KM	lymphoma; leukaemia; breast cancer; cirrhosis; neurological disorder;		
KM	stroke; Alzheimer's disease; multiple sclerosis; mental retardation;		
KM	cardiovascular disease; atherosclerosis; angina pectoris; indigestion;		
KM	congestive heart failure; gastrointestinal disorder; dysphagia; AIDS;		
KM	gastritis; autoimmune disorder; inflammatory disorder; Crohn's disease;		
KM	systemic lupus erythematosus; metabolic disorder; diabetes; obesity;		
KM	viral infection; herpesvirus; parvovirus;		
KW	acquired immune deficiency syndrome; ss.		
XX			
OS	Homo sapiens.		
PN	WO200190359-A2.		
XX			
PD	29-NOV-2001.		
XX			
XX	22-MAY-2001; 2001WO-US16833.		
XX			
PR	22-MAY-2000; 2000US-206222P.		
PR	25-MAY-2000; 2000US-207476P.		
PR	02-JUN-2000; 2000US-208834P.		
PR	02-JUN-2000; 2000US-208861P.		
PR	07-JUN-2000; 2000US-209868P.		
XX			
PA	(INCY-) INCYTE GENOMICS INC.		
PI	Patterson C, Tribouley CM, Yao MG, Griffin JA, Thornton M, Lu Y;		
PI	Kallick Da, Gandhi AR, Au-Young J;		
XX			
DR	WPI: 2002-106199/14.		
XX	P-PSDB; AA080493.		
XX			
PT	New G-protein coupled receptors useful for treating or preventing cell		
PT	proliferative (e.g. leukemia), neurological (e.g. stroke),		
PT	cardiovascular or autoimmune/inflammatory disorders		
XX			
PS	Claim 5; Page 137-138; 148pp; English.		
XX			
CC	The invention relates to a novel human G-coupled receptor (I). (I) and		
CC	its corresponding polynucleotides are useful for diagnosing, treating or		
CC	preventing cell proliferative diseases (e.g. lymphoma, leukaemia, breast		
CC	cancer or cirrhosis), neurological disorders (e.g. stroke, Alzheimer's		
CC	disease, multiple sclerosis or mental retardation), cardiovascular		
CC	diseases (e.g. atherosclerosis, angina pectoris or congestive heart		
CC	failure), gastrointestinal disorders (e.g. dysphagia, indigestion or		
CC	gastritis), autoimmune/inflammatory disorders (e.g. AIDS, Crohn's disease		
CC	or systemic lupus erythematosus) or metabolic disorders (e.g. diabetes		
CC	or obesity), or viral infections (e.g. infection by herpesvirus or		
CC	parvovirus). ABK1615-ABK1637 represent novel human G-coupled		
CC	receptor coding sequences of the invention.		

xx	Sequence	2739 BP	415 A	951 C	872 G	501 T	0 other:
Query Match	81.7%;	Score	2090.6;	DB	24;	Length	2739;
Best Local Similarity	91.4%;	Pred.	No. 0;				
Matches	2455;	Conservative	0;	Mismatches	19;	Indels	213;
						Gaps	14;
QY	1	ATGCGGGGCGCCGCGTCTCTGGGCGCCGACCCCTGCGGCGCTGCGTGCACACCTGGGAGCGGG	60				
DB	1	ATGCGGGGCGCCGCGTCTCTGGGCGCCGACCCCTGCGGCGCTGCGTGCACACCTGGGAGCGGG	60				
QY	61	GCCCATTTGTGCTCTTCACAGCACTTATGAGTAAGGAGGGGACTACGTGCGGGGGCGTG	120				
DB	61	GCCCATTTGTGCTCTTCACAGCACTTATGAGTAAGGAGGGGACTACGTGCGGGGGCGTG	120				
QY	121	TTCCCCCTTGGGCGAGAGCGGAGAGGCTGGGCTCCGACGCCGACAGCGGCCACAGCCCT	180				
DB	121	TTCCCCCTTGGGCGAGAGCGGAGAGGCTGGGCTCCGACGCCGACAGCGGCCACAGCCCT	180				
QY	181	GGTGTCACACAGGTTTTCCTCAACAGCGGCGCTGTGGGCACTGGGCATGAAATGGCCGTG	240				
DB	181	GGTGTCACACAGGTTTTCCTCAACAGCGGCGCTGTGGGCACTGGGCATGAAATGGCCGTG	240				
QY	241	GAGGAGATCAACAAACAAGTCGATGCTGCTGCCGGGCTGCGCTGGGCTACGACTCTT	300				
DB	241	GAGGAGATCAACAAACAAGTCGATGCTGCTGCCGGGCTGCGCTGGGCTACGACTCTT	300				
QY	301	GATACGTGCTCGAGACCTGTGTGGCCATGAGAGCCAGGCTCATGTCTTGCCCAAGGCA	360				
DB	301	GATACGTGCTCGAGACCTGTGTGGCCATGAGAGCCAGGCTCATGTCTTGCCCAAGGCA	360				
QY	361	GGCAGCCCGGACATGCGCGCTTACTGCAACTACACGCACTACAGCCCGGTGTGCTGCT	420				
DB	361	GGCAGCCCGGACATGCGCGCTTACTGCAACTACACGCACTACAGCCCGGTGTGCTGCT	420				
QY	421	GTGATCGGGGCGCCCATCGTCGACAGGTGCGCATAGGTCACCGGAGGGAAGTCTTGACGCTTTC	480				
DB	421	GTGATCGGGGCGCCCATCGTCGACAGGTGCGCATAGGTCACCGGAGGGAAGTCTTGACGCTTTC	480				
QY	481	CTCATGCCCC-----	489				
DB	481	CTCATGCCCC-----	489				
QY	490	-----CAGGTGAGGTAGGTCGTAGCATGAG	516				
DB	541	GCCCTGTGTCAAGAGATGCTCTTGGCCCTTGCAAGGTAGCTAGCGTCTAGCATGAG	600				
QY	517	CTGCTGAGCGCCCGGGAGACCTTCCGCTCTTCTTCGCAACCGTGCACGACACCGTGTG	576				
DB	601	CTGCTGAGCGCCCGGGAGACCTTCCGCTCTTCTTCGCAACCGTGCACGACACCGTGTG	660				
QY	577	CAGCTGACGCGCCCGGGAGCTGCTGCAAGAGTTGCGCTGGAACCTGGGTGGCCGCTTG	636				
DB	661	CAGCTGACGCGCCCGGGAGCTGCTGCAAGAGTTGCGCTGGAACCTGGGTGGCCGCTTG	720				
QY	637	GGCAGCGAGCAAGATAGCGCGCGGAGGAGCCGTGACATCTTTCGAGCCCTGGCCGCGGA	696				
DB	721	GGCAGCGAGCAAGATAGCGCGCGGAGGAGCCGTGACATCTTTCGAGCCCTGG-CTCGGCA	779				
QY	697	GCGGCGATCTGCAATGCGCAACAGAGGAGCTGTGTCGCTGCGCCGTCGCGATACTCGCGG	756				
DB	780	GCGGCGATCTGCAATGCGCAACAGAGGAGCTGTGTCGCTGCGCCGTCGCGATACTCGCGG	822				
QY	757	CTGGGGAGGATGTCAGAGAGTCTGTGACCAAGTGAACACAGAGAGGCTGAGTGGTGTG	816				
DB	823	-----GTGCAAGAGATCTGTGACCAAGTGAACACAGAGAGGCTGAGTGGTGTG	873				
QY	817	CTGTTCGCGCTCGTGTACAGCGCGCCACGACCCCTTCACTACAGATAGCAGCAGGCTC	876				
DB	874	CTGTTCGCGCTCGTGTACAGCGCGCCACGACCCCTTCACTACAGATAGCAGCAGGCTC	933				
QY	877	TCGCCCAAGGTGTGGGTGCGCAAGCGAGGCTTGCTGACCTTGACCTGAGTGTGCGGCTG	936				

Db 934 TCGCCAAGTGTGGTGCCAGCAGGCGCTGCTGACCTTGACCTGGTGTATGGGGCTG 993
 QY 937 CCGGCAATGGCCAGATGGGCGACGGTGTGCTTCTCTCAGAGGGGTGCCAGCTGCAC 996
 Db 994 CCGGCAATGGCCAGATGGGCGACGGTGTGCTTCTCTCAGAGGGGTGCCAGCTGCAC 1053
 QY 997 GAGTTCGCCCACTAGCTGAGAGCAGCAGCAGGCGCTGGCCACCGAGCCGGCTTCTGCT 1056
 Db 1054 GAGTTCGCCCACTAGCTGAGAGCAGCAGCAGGCGCTGGCCACCGAGCCGGCTTCTGCT 1113
 QY 1057 GCCCTGGGCGAGAGGAGAGGGGTGTGAGAGGAGCAGTGTGTGGCCAGCGCTGCCGAG 1116
 Db 1114 GCCCTGGGCGAGAGGAGAGGGGTGTGAGAGGAGCAGTGTGTGGCCAGCGCTGCCGAG 1173
 QY 1117 TGTGACTGATCAGCTGTGAGAGCAGTGAAGCGAGGGCTTAATCACCACGACGTTCT 1176
 Db 1174 TGTGACTGATCAGCTGTGAGAGCAGTGAAGCGAGGGCTTAATCACCACGACGTTCT 1196
 QY 1177 GTCACGACGCTGTGTATGAGGTGGCCAGGCGCTGCACAACTTTCAGTGAACGCGC 1236
 Db 1197 -----CGTCCCAAGGCGCTGCACAACTTTCAGTGAACGCGC 1236
 QY 1237 TCAGGCTCCCGCGCAGAGACCCCTGTAAGCCCTGGCAGCTCTGAGAACATGTACAAC 1296
 Db 1237 TCAGGCTCCCGCGCAGAGACCCCTGTAAGCCCTGGCAGCTCTGAGAACATGTACAAC 1296
 QY 1297 CTGACCTTCACAGTGGGGGGGTGTGCGGCTGCGGTTCGACAGAGGGGAAAGCTGACATG 1356
 Db 1297 CTGACCTTCACAGTGGGGGGGTGTGCGGCTGCGGTTCGACAGAGGGGAAAGCTGACATG 1356
 QY 1357 GAGTACGACCTGTAAGCTGTGGGTGTGAGAGGGGTGAGTCCAGGCTCCACAGACGTGGGC 1416
 Db 1357 GAGTACGACCTGTAAGCTGTGGGTGTGAGAGGGGTGAGTCCAGGCTCCACAGACGTGGGC 1416
 QY 1417 AGGTTCAAACGGCAGCTTCAAGACAGAGCGCTGAAGATCCGTGCACACGTCTGAACAC 1476
 Db 1417 AGGTTCAAACGGCAGCTTCAAGACAGAGCGCTGAAGATCCGTGCACACGTCTGAACAC 1476
 QY 1477 -----CAGAGCCCGTGTCCCGGGTCCGCGG 1503
 Db 1477 CAGCCGACGACGACGACGACGACGCGCTGTGTGCGAGAAAGCCGCTGTCCGGGTGCGG 1536
 QY 1504 CAGTCCAGAGAGGCGCAGGTGCGCGGGTCAAGGGGTTCACATCTGCTGACGACTGT 1563
 Db 1537 CAGTCCAGAGAGGCGCAGGTGCGCGGGTCAAGGGGTTCACATCTGCTGACGACTGT 1596
 QY 1564 GTGGACTCGAGGCGGCGGACGTACCGCAAAACCCAGACGACATGCGCTGACCTTTGT 1623
 Db 1597 GTGGACTCGAGGCGGCGGACGTACCGCAAAACCCAGACGACATGCGCTGACCTTTGT 1656
 QY 1624 GGCCAGAGATGAGTGTCCCGGAGGAGGAGACAGCTGCTCCGCGAGGCTGTGGTTC 1683
 Db 1657 GGCCAGAGATGAGTGTCCCGGAGGAGGAGACAGCTGCTCCGCGAGGCTGTGGTTC 1716
 QY 1684 CTGGCATGGGGGAGCAGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1743
 Db 1717 CTGGCATGGGGGAGCAGCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1776
 QY 1744 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1803
 Db 1777 CTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1836
 QY 1804 TCGGGGGGGGGCGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1863
 Db 1837 TCGGGGGGGGGCGCTGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1896
 QY 1864 CTCTGTTTCCCTGAGCAGCCAGCCCTGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1923
 Db 1897 CTCTGTTTCCCTGAGCAGCCAGCCCTGAGCATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1956
 QY 1924 CTCCCGCTCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1983
 Db 1957 CTCCCGCTCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2016

QY 1984 TCAGAACTCCCTCTGAGCTGGGCGACAGCCGCTGAGTGTGCTGCGGGGGCCCTGGGCGC 2043
 Db 2017 TCAGAACTCCCTCTGAGCTGGGCGACAGCCGCTGAGTGTGCTGCTGCGGGGGCCCTGGGCGC 2076
 QY 2044 TGGCTGTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2103
 Db 2077 TGGCTGTGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2136
 QY 2104 GCCCTCCCGCGAGAGTGTGAC - GACTGAGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2162
 Db 2137 GCCCTCCCGCGAGAGTGTGACTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2196
 QY 2169 CTGCGGACACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2222
 Db 2197 CTGCGGACACGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2255
 QY 2223 CTTCCTGCTTC - CTGGGCACTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2281
 Db 2256 CTTCCTGCTTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2315
 QY 2282 CC - - CGTGGCCTCACTTTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2336
 Db 2316 CCAGCTGGCTCAGACATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2375
 QY 2337 GCCCTCCCTGCG - CANTGTGAGGTGT - - CCTCAGCGCCCGCGCTGCTGCTGCTGCTGCTGCT 2393
 Db 2376 GCCCTCCCTGCGACATGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2435
 QY 2394 CCTGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2453
 Db 2436 CCTGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2495
 QY 2454 GCGGACGCCAGGCTCAACACCCCGAGTTCCTTCTGAGGGGGCCCTGGGAGTGC - C 2512
 Db 2496 GCGGACGCCAGGCTCAACACCCCGAGTTCCTTCTGAGGGGGGGCCCTGGGAGTGCAC 2555
 QY 2513 AAGGCGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2559
 Db 2556 AAGGCGAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2602

RESULT 8
 ID ALU38463 standard; DNA; 8001 BP.
 AC ALU38463;
 XX 15-AUG-2002 (first entry)
 DE XX
 XX Genomic DNA encoding a human G-protein coupled receptor (GPCR).
 KW G-protein coupled receptor; GPCR; gene chip; human; immune response;
 KW chromosomal position; transgenic animal; gene therapy; gene; ds..
 OS Homo sapiens.
 FH key
 FT CDS
 FT 2118..5236
 FT location/qualifiers
 FT /tag= a
 FT /product= "Human G-protein coupled receptor"
 FT /note= "this coding sequence contains 5 introns"
 FT 2118..2308
 FT /tag= b
 FT /number= 1
 FT 2309..2409
 FT /tag= c
 FT /number= 1
 FT 2410..2710
 FT /tag= d
 FT /number= 2
 FT 2711..2795
 FT /tag= e
 FT intron

Db 3018 CACGAGGGGCTGTGTCGCGTCCGCTGACGATGACTGCGGCTGGGAGAGGTGAGGAC 3077
Qy 775 GTCTGCAACAGGTGAACAGAGACAGCGTGAAGTGTGTCTGTCTGCTCGCTGAC 834
Db 3078 GTCTGCAACAGGTGAACAGAGACAGCGTGAAGTGTGTCTGTCTGCTCGCTGAC 3137
Qy 835 GCGGCGCCAGGCGCTCTCACTAGCAATGACAGACAGGCTGTGCGCAAGGTGTGGGTG 894
Db 3138 GCGGCGCCAGGCGCTCTCACTAGCAATGACAGACAGGCTGTGCGCAAGGTGTGGGTG 3197
Qy 895 GCCACGAGGCGCTGTGCTGACCTGTGATGAGGAGTGTGCGGATGGCGCAGATG 954
Db 3198 GCCACGAGGCGCTGTGCTGACCTGTGATGAGGAGTGTGCGGATGGCGCAGATG 3257
Qy 955 GGCAGAGGTGTGGTGTCTCTCCAGAGGGGTGCCAGCTGACAGATTTCCCAAGTACGTG 1014
Db 3258 GGCAGAGGTGTGGTGTCTCTCCAGAGGGGTGCCAGCTGACAGATTTCCCAAGTACGTG 3317
Qy 1015 AAGACGCACTGTGGCCCTGGCCACGACCGGCTTCTGTCTGCGCTGGGAGAGAGGAG 1074
Db 3318 AAGACGCACTGTGGCCCTGGCCACGACCGGCTTCTGTCTGCGCTGGGAGAGAGGAG 3377
Qy 1075 CAGGCTGTGAGAGGAGAGAGTGTGGGCGAGCGCTGCCAGTGTGACTGATCAGCTG 1134
Db 3378 CAGGCTGTGAGAGGAGAGAGTGTGGGCGAGCGCTGCCAGTGTGACTGATCAGCTG 3437
Qy 1135 CAGAACGTGAGAGCGAGGCTTAATACCAACAGAGCTTCTGTCTGAGCAGCTGTAT 1194
Db 3438 CAGAACGTGAGAGCGAGGCTTAATACCAACAGAGCTTCTGTCTGAGCAGCTGTAT 3497
Qy 1195 AGCGTGGCCAGAGGCGCTGACAACTGTCTAGTGAAGCGCTTCAAGGCTGCCCGCGAG 1254
Db 3498 AGCGTGGCCAGAGGCGCTGACAACTGTCTAGTGAAGCGCTTCAAGGCTGCCCGCGAG 3557
Qy 1255 GACCCCGTGAAGCCCTG----- 1271
Db 3558 GACCCCGTGAAGCCCTG----- 1271
Qy 1272 ----- 1271
Db 3618 GCCCAGGCCACGAGGACGCGCCACGACGCTGAGCTGAGGTGGCTGGCGGCTCAGCCCC 3677
Qy 1272 -----GACACTCTCTGAGAACTGTACAACTGACCTTCCAGCTGGCGGGCTG 1320
Db 3678 GTCGCCCGCGCGACCTCTGTGAGAACTGTACAACTGACCTTCCAGCTGGCGGGCTG 3737
Qy 1321 CCGCTGGCGGTTCGACAGCAGCGGAAAGTGTGACATGAGATGACCTGAGCTGTGGTG 1380
Db 3738 CCGCTGGCGGTTCGACAGCAGCGGAAAGTGTGACATGAGATGACCTGAGCTGTGGTG 3797
Qy 1381 TGGCAGGCGTCAGTGCGCAGGCTTCAGCAGCGTGGGAGGTTAAAGGAGCCTCAGAGCA 1440
Db 3798 TGGCAGGCGTCAGTGCGCAGGCTTCAGCAGCGTGGGAGGTTAAAGGAGCCTCAGAGCA 3857
Qy 1441 GAGCGCTGAAGATCCGCTGGCACAAGTGTGACAAC----- 1476
Db 3858 GAGCGCTGAAGATCCGCTGGCACAAGTGTGACAACAGGTGAGGTGAGGTGGGTGTGC 3917
Qy 1477 ----- 1476
Db 3918 CAGGCGTCCCGTGTAGCCCCCGCGCAGGGCGCAGCCTGTGGGGTGGGGCCGTTCCAG 3977
Qy 1477 -----CAGAAAGCCG 1486
Db 3978 TCTCCCGTGGGCAATGCCAGCGAGACAGACCCAGGCGCTGTGTGGCAAGAGCCG 4037
Qy 1487 TGTCCCGGTGTCTGCGGAGAGTGTCCAGAGAGGCGCAGTGGCGCGGTTCAAGGGTTCCACT 1546
Db 4038 TGTCCCGGTGTCTGCGGAGAGTGTCCAGAGAGGCGCAGTGGCGCGGTTCAAGGGTTCCACT 4097
Qy 1547 CCGTGTGCTAGCAGTGTGTGTGAGCTGAGGCGGCGCAGCTACCGGCAAAAC----- 1597

Db 4098 CCTGTGCTAGCAGTGTGTGAGACTGAGAGGCGGCGAGTACCGGCAAAACCCAGGTGAGC 4157
Qy 1598 ----- 1597
Db 4158 GCGCTTCCCGGAGCGCGGGGTGGGAGCAGACAGGGAGGTCTGTCCAACTCTGACT 4217
Qy 1598 -----CAG 1600
Db 4218 CTGAGACCAAGAGCCCAAGAGGAGCAAGAGCAACACCAGCGCCCTTCTCTCTCAGAG 4277
Qy 1601 ACGACATGCGCTGCACTTTTGTGGCCAGATGAGTGTCTCCCGGAGCGAAGCAGACGCT 1660
Db 4278 ACGACATGCGCTGCACTTTTGTGGCCAGATGAGTGTCTCCCGGAGCGAAGCAGACGCT 4337
Qy 1661 GCTTCCCGGAGGTCTGCGTGTCTGGCATGGGGGAGCGCGGCTGTGTCTGTCTGTCC 1720
Db 4338 GCTTCCCGGAGGTCTGCGTGTCTGGCATGGGGGAGCGCGGCTGTGTCTGTCTGTCC 4397
Qy 1721 TGTCTGTGAGCTGTGGCGGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1780
Db 4398 TGTCTGTGAGCTGTGGCGGTGGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4457
Qy 1781 GGGACAGCCCACTGTGTAGAGCTGTGGGGGGGCGCCCTGGCTGTGTGTGTGTGTGTGTGT 1840
Db 4458 GGGACAGCCCACTGTGTAGAGCTGTGGGGGGGCGCCCTGGCTGTGTGTGTGTGTGTGTGT 4517
Qy 1841 TGGGCGT 1900
Db 4518 TGGGCGT 4577
Qy 1901 TGGCCAGCAGCGCTTGT 1960
Db 4578 TGGCCAGCAGCGCTTGT 4637
Qy 1961 AGCGGCGAGATCTTGT 2020
Db 4638 AGCGGCGAGATCTTGT 4697
Qy 2021 GCTGCTGTGGGGGCGCTGTGGGCGTGTGGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2080
Db 4698 GCTGCTGTGGGGGCGCTGTGGGCGTGTGGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4757
Qy 2081 CACTGTGCACTGT 2140
Db 4758 CACTGTGCACTGT 4817
Qy 2141 TGGCCAGGAGGCGCTGT 2200
Db 4818 TGGCCAGGAGGCGCTGT 4877
Qy 2201 ACGCCACCAATGCAAGCGT 2260
Db 4878 ACGCCACCAATGCAAGCGT 4937
Qy 2261 AGCGGCGGT 2320
Db 4938 AGCGGCGGT 4997
Qy 2321 CCTGGGTCTCTTGT 2380
Db 4998 CCTGGGTCTCTTGT 5057
Qy 2381 AGATGGGCGCGCTCTGT 2440
Db 5058 AGATGGGCGCGCTCTGT 5117
Qy 2441 GTTACTGTCTATGTGGGAGCGCAGGAGCTCAACACCCGAGTTCCTGTGGAGGGGGCC 2500
Db 5118 GTTACTGTCTATGTGGGAGCGCAGGAGCTCAACACCCGAGTTCCTGTGGAGGGGGCC 5177
Qy 2501 CTGGGAGTGTCCCAAGGCGCAAGATGACGGGAGCAACAGGAATAGGGGAAACATGAGTGA 2559
Db 5178 CTGGGAGTGTCCCAAGGCGCAAGATGACGGGAGCAACAGGAATAGGGGAAACATGAGTGA 5236


```

Db 2911 ACCGTGACCTGCTGGCTTCCCGGAGGTGTGACGGATGCGACATCTGCTCCACG 2970
QY 2149 GAGCGCTGATGACACTGCGGACACAGGCTCTGGGTAGCTGGGCTAGGCGACGCGAC 2208
Db 2971 GAGCGCTGATGACACTGCGGACACAGGCTCTGGGTAGCTGGGCTAGGCGACGCGAC 3030
QY 2209 AATGCCAAGCTGAGCTTCTCTGCTTCTGCTGAGGCACTTCTGCTGAGGCGAGCGAGG 2268
Db 3031 AATGCCAAGCTGAGCTTCTCTGCTTCTGCTGAGGCACTTCTGCTGAGGCGAGCGAGG 3090
QY 2269 TGTACACACCTGCGCGCTGCGCTACCTTTCATGCTGGCTACTTATACCTGGGCTC 2328
Db 3091 TGTACACACCTGCGCGCTGCGCTACCTTTCATGCTGGCTACTTATACCTGGGCTC 3150
QY 2329 TCTCTTGGCCCTCTCTGCGCAATGTGAGTGTCTCTGAGGCGCGCGCTGAGTGGGC 2388
Db 3151 TCTCTTGGCCCTCTCTGCGCAATGTGAGTGTCTCTGAGGCGCGCGCTGAGTGGGC 3210
QY 2389 GCGCTCTGCTGTGTCTGCTGAGGCACTGCTGCTTCTGCTGAGGCGAGGCTTACCTG 2448
Db 3211 GCGCTCTGCTGTGTCTGCTGAGGCACTGCTGCTTCTGCTGAGGCGAGGCTTACCTG 3270
QY 2449 CTGATGGGCGAGCGAGGCTTAACACCCCGAGTTCCTCTGAGGAGGCGCGCTGGGAT 2508
Db 3271 CTGATGGGCGAGCGAGGCTTAACACCCCGAGTTCCTCTGAGGAGGCGCGCTGGGAT 3330
QY 2509 GCGCAAGGCGCAATGACGAGGACAGCAAGAAATCAGGGAATCAGAGTGA 2559
Db 3331 GCGCAAGGCGCAATGACGAGGACAGCAAGAAATCAGGGAATCAGAGTGA 3381

RESULT 10
AADI7508
ID AADI7508 standard; DNA: 2687 BP.
XX
AC AADI7508;
XX
DT 10-DEC-2001 (first entry)
XX
DE Human taste receptor, hTIR3 genomic DNA - 3' contig.
XX
KW Human: taste-cell-specific G protein-coupled receptor; hTIR3; drug;
KW genetic modulation; pharmaceutical; taste sensation; food industry;
KW chemosensory transduction; chromosome 1p36.2-1p36.33; 68.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Intron 1..69
FT /*tag= a
FT /number= 1
FT /cons_splice= (5'site:NO, 3'site:YES)
FT /partial
FT 70..847
FT /*tag= b
FT /number= 1
FT Intron 848..961
FT /*tag= c
FT /number= 2
FT 962..1165
FT /*tag= d
FT /number= 2
FT Intron 1166..1299
FT /*tag= e
FT /number= 3
FT 1300..1420
FT /*tag= f
FT /number= 3
FT Intron 1421..1546
FT /*tag= g
FT /number= 4
FT 1547..2505

```

```

FT /*tag= h
FT /number= 4
FT Intron 2506..2687
FT /*tag= i
FT /number= 5
FT /cons_splice= (5'site:NO, 3'site:NO)
XX
PN W020016563-A2.
XX
PD 13-SEP-2001.
XX
PF 07-MAR-2001: 2001MO-US07265.
XX
PR 07-MAR-2000: 2000US-0187546.
PR 07-APR-2000: 2000US-0195536.
PR 06-JUN-2000: 2000US-0209840.
PR 23-JUN-2000: 2000US-0214213.
PR 17-AUG-2000: 2000US-0226448.
PR 03-JAN-2001: 2001US-0259227.
XX
PA (SENO-) SENOMYX INC.
XX
PI Adler JE, Zozulya S, O'Connell SM, Li X, Staszewski L;
XX WPI: 2001-582267/65.
XX
DR New mammalian taste-cell-specific G protein-coupled receptor
XX PT polypeptides for identifying compounds that modulate taste signaling
XX PT are useful in food, to modulate the sweet taste of foods or drugs -
XX
PS Claim 133; Page 69-71; 119pp; English.
XX
XX The invention relates to mammalian taste-cell-specific G protein-coupled
XX receptors, TIR and their corresponding cDNA molecules. Taste receptors,
XX TIR are useful for screening compounds which are used to activate or
XX modulate chemosensory transduction, such as taste sensation. The
XX identification and isolation of novel taste receptors and taste
XX signalling molecules allow for new methods of chemical and genetic
XX modulation of taste transduction pathways. The taste modulating
XX compounds are useful in pharmaceuticals and food industries to improve
XX the taste of a variety of consumer products, or to block undesirable
XX tastes, e.g., in certain pharmaceuticals. TIRs are also useful in
XX biochemical assay for identifying tastant (TIR) ligands having binding
XX specificity for TIR involved in taste signalling. The present sequence is
XX human taste-cell-specific G protein-coupled receptor, hTIR3 genomic DNA -
XX 3' contig. Human TIR3 gene is localised on chromosome 1p36.2-1p36.33.
XX
SQ Sequence 2687 BP; 403 A; 920 C; 885 G; 479 T; 0 other;
XX
Query Match 65.18; Score 1666; DB 22; Length 2687;
Best Local Similarity 84.7%; Pred. No. 7.9e-274;
Matches 2070; Conservative 0; Mismatches 0; Indels 374; Gaps 3;
QY 490 CAGGTGACCTAGGCTGTAGATGAGTGTGAGGCGCGGAGACCTCCCTCTTC 549
Db 62 CAGGTGACCTAGGCTGTAGATGAGTGTGAGGCGCGGAGACCTCCCTCTTC 121
QY 550 TTCCGACCGTGCACGACGAGCGTGTGACGAGCGCGCGGAGCTGTGACAGAG 609
Db 122 TTCCGACCGTGCACGACGAGCGTGTGACGAGCGCGCGGAGCTGTGACAGAG 181
QY 610 TTGGCTGGAACCTGGTGGCGCGCTGTGGCAGCAGCAGAGTACGGCCGAGAGGCTG 669
Db 182 TTGGCTGGAACCTGGTGGCGCGCTGTGGCAGCAGCAGAGTACGGCCGAGAGGCTG 241
QY 670 AGCATCTTCTGCGCGCTGCGCGGCGGACGAGTGTGATGCGGACGAGAGGCTGGTG 729
Db 242 AGCATCTTCTGCGCGCTGCGCGGCGGACGAGTGTGATGCGGACGAGAGGCTGGTG 301
QY 730 CCGCTGCGCGGCGGACGAGTGTGCGGAGGAGTGTGAGAGCGTCTGACACAGGTG 789
Db 302 CCGCTGCGCGGCGGACGAGTGTGCGGAGGAGTGTGAGAGCGTCTGACACAGGTG 361

```


ID	AD34459	standard; CDNA: 1788 BP.
XX	AD34459;	
XX	16-JUL-2002	(first entry)
XX	Human GPCR cDNA #2.	
XX	Human GPCR cDNA #2.	
XX	Human; G protein-coupled receptor; GPCR; central nervous system disease;	
XX	Cancer; anorexia; bulimia; asthma; benign prostatic hypertrophy; ulcer;	
XX	allergy; acute heart failure; myocardial infarction; urinary retention;	
XX	Parkinson's disease; Alzheimer's disease; hypotension; angina pectoris;	
XX	hypertension; osteoporosis; inflammation; infection; dementia; diabetes;	
XX	leukemia; lymphoma; gene therapy; gene; ss.	
XX	Homo sapiens.	
XX	Key	Location/Qualifiers
XX	CDS	1..1788
XX	FT	/*tag- a
XX	FT	/product= "Human GPCR protein #2"
XX	FT	/note= "CDS does not include stop codon"
XX	FT	/partial
XX	PN	WO200224885-A2.
XX	PD	28-MAR-2002.
XX	PP	21-SEP-2001; 2001WO-EP10929.
XX	PR	25-SEP-2000; 2000US-234982P.
XX	PA	(FARB) BAYER AG.
XX	PI	Kossida S;
XX	DR	WPI: 2002-339949/37.
XX	DR	P-PSDB; AAE21818.
XX	PT	New G-protein coupled receptor polynucleotide and its encoded protein,
XX	PT	useful for identifying modulators of GPCR activity, and in gene therapy
XX	PT	for treating bacterial infection, cancer, acute heart failure or
XX	PT	inflammation
XX	PS	Claim 20; Fig 4; 101pp; English.
XX	CC	The invention relates to human G protein-coupled receptor (GPCR)
XX	CC	polypeptides and polynucleotides. GPCR sequences are useful for
XX	CC	preventing or ameliorating the diseases such as cancer, anorexia,
XX	CC	bulimia, asthma or other allergies, peripheral or central nervous
XX	CC	system disease, acute heart failure, hypotension, hypertension,
XX	CC	urinary retention, osteoporosis, angina pectoris, myocardial
XX	CC	infarction, ulcer, inflammation or benign prostatic hypertrophy
XX	CC	and infections (e.g. bacterial, fungal, protozoan or viral).
XX	CC	Other diseases that may be treated include Alzheimer's disease,
XX	CC	Parkinson's disease, dementia, diabetes, leukemia and lymphoma.
XX	CC	Sequences of the invention are used for identifying agonists and
XX	CC	antagonists that modulate their activity. They are also used in
XX	CC	gene therapy. The present sequence is human GPCR cDNA.
XX	XX	Sequence 1788 BP; 236 A; 620 C; 593 G; 339 T; 0 other;
XX	SO	
XX	Query Match	61.5%; Score 1574; DB 24; Length 1788;
XX	Best Local Similarity	89.8%; Pred. NO. 3.1e-25;
XX	Matches 1788; Conservative	0; Mismatches 0; Indels 204; Gaps 1;
XX	490 CAGGTACAGTACGGTGTAGCATGAGAGTGTGACGCGCCGGAGACCTTCCCTCCTTC	549
XX	1 CAGGTACAGTACGGTGTAGCATGAGTGTGACGCGCCGGAGACCTTCCCTCCTTC	60
XX	550 TTCGACACCTGTCCACGAGACCTGTGACGCTGACGCGCCGGAGCTGCTGACGAG	609

Db	61	TTTGGAACCGTGGCCACAGCAACGCTGTGTGCACTGACGGCCGGCGGAGTGTCTGTGAGGAG	120
Qy	610	TTTGGCTGGAACTGGGTGTGGCCGCTTGTGGCAGCGACGAGTACGGCCGGCAGGGCCTG	669
Db	121	TTTCGGCTGGAACTGGGTGTGGCCGCTTGTGGCAGCGAGTACGAGTACGGCCGGCAGGGCCTG	180
Qy	670	AGCATCTTCTCGGCCCTTGGGCGGGGGGCAACGGGCAATCTGCATGGCCGACAGAGGGCCTGGTG	729
Db	181	AGCATCTTCTCGGCCCTTGGGCGGGGGGCAACGGGCAATCTGCATGGCCGACAGAGGGCCTGGTG	240
Qy	730	CCGCTGGCCCCGTGGCCGATGATCGCGGCTTGGGCAAGGTGACAGAGCTCTCTGACCAAGGTG	789
Db	241	CCGCTGGCCCCGTGGCCGATGATCGCGGCTTGGGCAAGGTGACAGAGCTCTCTGACCAAGGTG	300
Qy	790	AACCAAGACACAGCTGTGACAGTGTGTCTGTCTGCTCCGTGTGACAGCGCCGACGCTCTC	849
Db	301	AACCAAGACACAGCTGTGACAGTGTGTCTGTCTGCTCCGTGTGACAGCGCCGACGCTCTC	360
Qy	850	TTTCAACTACAGCAATCACACACAGGCTCTGCCCCAAGTGTGGGTGGCCAGCGAGGCGCTGG	909
Db	361	TTTCAACTACAGCAATCACACACAGGCTCTGCCCCAAGTGTGGGTGGCCAGCGAGGCGCTGG	420
Qy	910	CTGACCTCTGACCTGTGTCAATGGGGGCTGCCGGGCAATGGCCAGATGGGCAAGGCTGTGGCC	969
Db	421	CTGACCTCTGACCTGTGTCAATGGGGGCTGCCGGGCAATGGCCAGATGGGCAAGGCTGTGGCC	480
Qy	970	TTTCTCTCAGAGGGGTGCCACAGCTGTGACAGAGTTCCCCAGTACGTGTGAAGCGACCTGTGCC	1029
Db	481	TTTCTCTCAGAGGGGTGCCACAGCTGTGACAGAGTTCCCCAGTACGTGTGAAGCGACCTGTGCC	540
Qy	1030	CTGGGCACACGACCCGGGCTTCTGCTGCTGCCCCGTGGGGAGAGGGAGCAGAGGCTGTGAGAG	1089
Db	541	CTGGGCACACGACCCGGGCTTCTGCTGCTGCCCCGTGGGGAGAGGGAGCAGAGGCTVTGAGAGAG	600
Qy	1090	GACCTGGTGGGGCAGCGCTCTGCCGACAGTGTGACATGACCTGTGAGAACCTGTGACGCA	1149
Db	601	GACCTGGTGGGGCAGCGCTCTGCCGACAGTGTGACATGACCTGTGAGAACCTGTGACGCA	660
Qy	1150	GGGCTAAATCACACACAGAGCGTTCTGTCTACGACAGCTGTGTATAGCGTGGCCGACGCC	1209
Db	661	GGGCTAAATCACACACAGAGCGTTCTGTCTACGACAGCTGTGTATAGCGTGGCCGACGCC	720
Qy	1210	CTGCAACAACACTCTTCAGTGTCAACGGCTCAGAGTGGCCCGGCGACAGAACCCGTGAAAGCC	1266
Db	721	CTGCAACAACACTCTTCAGTGTCAACGGCTCAGAGTGGCCCGGCGACAGAACCCGTGAAAGCC	780
Qy	1270	TGCGAGCTCCTGGAGAACATGTACAACTGTACACTTCCAGTGGGGGGGCTGCGCGCTCGGG	1329
Db	781	TGCG-----	783
Qy	1330	TTTCAGACAGACAGCGGAAACGTGTGACATGTGAGATGACACTGAGCTGTGGGTGTGGCAGGGC	1389
Db	784	-----	783
Qy	1390	TCAATGCCCAAGGCTCCACAGACGTGGGGCAGTTCAACGGCAGCCTCAGAGACAGAGCCCTG	1449
Db	784	-----	783
Qy	1450	AAGATCCGCTGGGCACACAGCTGTGACAAACAGAAAGCCGTTGCCGGGCTCGCGGGCAGTGC	1509
Db	784	-----CAAGAACCCGTTGTCCCGGTTCTCGGGGAGAGTGC	816
Qy	1510	CAGAGAGGCGCAGGTCGCGCGGCTCAAGAGGGGTTTCACTCTGCTGTGTACGACTGTGTGAC	1569
Db	817	CAGAGAGGCGCAGGTCGCGCGGCTCAAGAGGGGTTTCACTCTGCTGTGTACGACTGTGTGAC	876
Qy	1570	TGCGAGGCGGGCAGCTTACCGGCAAAAACAGAGACATGCGCTGTGACCTTTTGTGGCCAG	1629
Db	877	TGCGAGGCGGGCAGCTTACCGGCAAAAACAGAGACATGCGCTGTGACCTTTTGTGGCCAG	936
Qy	1630	GATAGTGTGTCTCCCGAGAGCGAAGACACAGCTGTCTCCCGCGCAGTCTGTGTTCTTGCA	1688
Db	937	GATAGTGTGTCTCCCGAGAGCGAAGACACAGCTGTCTCCCGCGCAGTCTGTGTTCTTGCA	996

Db 2452 AATGCTATGACTTCTGTGGTGGCCAGAGCTCAACACCCAGAGATGCTCTGGAGG 2511
 Oy 2497 GGGCCGTGGGATGCCAGGCGCAGAATGACGGGA 2530
 Db 2512 AGCCCAAGAGAGCATGATGGATAGTGSTA 2545

RESULT 13
 AAS97393
 ID AAS97393 standard; cDNA: 2577 BP.
 AC AAS97393;
 XX 12-MAR-2002 (first entry)
 DT
 DE Murine SACL gene cDNA.
 XX Human; mouse; SACL; carbohydrate; sweetener; ethanol; alcoholism; ss;
 RW obesity; diabetes; transgenic embryo; body tissue; body fluid; pancreas;
 RW blood; tongue; PCR primer; anorectic; antidiabetic; gene therapy;
 KW protein replacement therapy.
 XX Mus sp.
 OS
 XX W0200183749-A2.
 PN
 XX 08-NOV-2001.
 PD
 XX 25-APR-2001; 2001WO-US13387.
 PE
 XX 28-APR-2000; 2000US-200794P.
 PR 28-JUL-2000; 2000US-221419P.
 PR 10-NOV-2000; 2000US-247443P.
 XX
 PA (WARN) WARNER LAMBERT CO.
 PA (NONE) MONELL CHEM SENSES CENT.
 PI Bachmanov AA, Beauchamp GK, Chatterjee A, De Jong PJ, Li S, Li X;
 PI Ohmen JD, Reed DR, Ross D, Tordoff MG;
 XX WPI: 2002-075162/10.
 DR P-PSDB; AAU73183.
 XX Novel isolated polypeptide comprising variant form of mouse or human
 PT SACL polypeptide, and is associated with altered preference for
 PT carbohydrates or other sweeteners, useful for preventing obesity,
 PT diabetes, alcoholism
 XX
 PS Claim 1; Page 28-29; 239pp; English.
 XX The invention relates to an isolated polypeptide, comprising a variant
 CC form of mouse or human SACL polypeptide. The variant form is associated
 CC with altered preference for carbohydrates, other sweeteners or ethanol.
 CC The polypeptide and its associated DNA sequence can be produced by
 CC recombinant techniques and is useful for preventing obesity, diabetes or
 CC alcoholism associated with SACL expression. The sequences are useful in
 CC screening for drugs and sweeteners. Recombinant cell lines and transgenic
 CC embryos may be used in screening for and identifying agents that induce
 CC or repress function of SACL. Predisposition to diabetes, obesity or
 CC alcoholism can be ascertained by testing any fluid or tissue of a human
 CC (such as blood, pancreas or tongue) for sequence variations of the SACL
 CC gene. A sequence variation of the SACL locus may indicate a
 CC predisposition to diabetes, obesity and/or alcoholism and may provide a
 CC diagnostic mark. The polynucleotide can be detected in a biological
 CC sample by contacting the DNA with a probe to form a hybridisation complex
 CC which is then detected. The sequences represent cDNA encoding human and
 CC mouse SACL polypeptides and PCR primers specific for the SACL genes.
 XX
 SQ Sequence 2577 BP; 519 A; 735 C; 689 G; 634 T; 0 other;

Query Match 56.9%; Score 1454.8; DB 24; Length 2577;
 Best Local Similarity 74.3%; Pred. No. 5.2e-238;
 Matches 1891; Conservative 0; Mismatches 622; Indels 33; Gaps 3;

Oy 1 ATGCTGGGCGCTGCTGCTGGGCTCAAGCCTCTGGGCTCTGACCTGGAGCGGG 60
 Db 1 ATGCCAGCTTTGGCTATATGAGTGTCTCAGCCGTGCTTCTGAGCTTGGGATGGGG 60
 Oy 61 GCGCCATGTGCTGTCTCAGCAGCACTAGATGAAGAGGGAGACTAGCTGGGGGGCTG 120
 Db 61 GCGCTTTGTGTGTCTCAGCAGCACTAGATGAAGAGGGAGACTAGCTGGGGGGCTA 120
 Oy 121 TTCCCTTGGGGAGGCGGAGAGGCTGGCTCGCAGCCGAGACAGGCCAGACGCTT 180
 Db 121 TTTCCCTGGGCTCAACCGAGAGGCGCACTCTCAACAGAGAACACCAACCAAGCATC 180
 Oy 181 GTGTGACAGGCTTCTCTCAACAGCGCTCTGGGAGCAGCGCCATTAATATGGCCGTG 240
 Db 181 CCGTCAACAGGTTCTCAACCCCTTGGTTGTCTGGGCGCATGGCTATCAAGATGGCTGTG 240
 Oy 241 GAGAGATCAACCAACCACTGATCTGCTGGCCGGGCTGGCGCTGGCTAGACCTCTTT 300
 Db 241 GAGAGATCAACCAACCACTGATCTGCTGGCCGGGCTGGCGCTGGCTAGACCTATTT 300
 Oy 301 GATACGCTCTGGAGCCTGTGTGGCCATGAAGCCAGCTCATTTCTTGGCCAGGCA 360
 Db 301 GACACATGCTCGAGCCAGCTGATGACATGAATCCAGTCTCATGTTCTGGCCAGGCTG 360
 Oy 361 GGCAGCCGCGACATGCGCGCTACTGCACTACAGCAGATCAAGCCCGGCTGGCTGCT 420
 Db 361 GGCAGTCAACAGCATGCTGCTGCTACTGCACTACAGCAGATCAACCCCGGCTGGCTGCT 420
 Oy 421 GTCATGCGGCCCCACTGCTCAGAGCTCGCATGATGTCACCGGCAAGTTCTTACCTTTC 480
 Db 421 GTCATGCGGCCCCACTGCTCAGAGCTCGCATGATGTCACCGGCAAGTTCTTACCTTTC 480
 Oy 481 CTCATGCCCCAGGTGATGCTAGCGTCTAGCATGAGCTGTGAGGCCCCGGAGACCTTC 540
 Db 481 CTCATGCCCCAGGTGATGCTAGCGTCTAGCATGAGCTGTGAGGCCCCGGAGACCTTC 540
 Oy 541 CCCCTCTCTTCCGACAGCTGCGCCAGCAGCGTGTGACAGTACAGCCGCGGAGCTG 600
 Db 541 CCATCTCTTCCGACAGTGTGCGCCAGTGTGACAGTGTGACAGTGTGACCTGTG 600
 Oy 601 CTGCAGAGATTCGCTGGAACCTGAGTGGCCGCGCTGGGAGCAGCAGATGACGCGCG 660
 Db 601 TTGCAGAACTTCAGTGAAGCTGAGTGGCCGCGCTTGAAGGAGATGATGACTATGACGCG 660
 Oy 661 CAGGCTGTGAGCATTTCTCGGCTGCGCGGAGCAGCGCATGTGATGCGGACAG 720
 Db 661 GAAGCTGTGAGCATTTCTGATGCTGCGCAATGACAGCAGGATGATGATGACATGAG 720
 Oy 721 GAGCTGTGCGGCTGCGCCGTCGATGACATGCGGCTGGGAGGATGACAGACGCTCTG 780
 Db 721 GAGCTGTGCGGCTGCGCCGTCGATGACATGCGGCTGGGAGGATGACAGACGCTCTG 780
 Oy 781 CACCAAGTGAACAGCAGCAGGCTGAGTGTGCTGTTCCTCCGCTGACAGCAGCC 840
 Db 781 CCGCAAGTGAACCAAGTGAAGTGAAGTGTGCTGTTCCTCCGCTGACAGCAGCC 840
 Oy 841 CAGCGCTCTTCACTAAGATGATGACAGCAGCTCTGCGCCAGATGATGATGATGATG 900
 Db 841 TACTCTCTTTTAACTAAGATGATGACAGCAGCTCTGCGCCAGATGATGATGATGATG 900
 Oy 901 GAGGCTGTGACCTGTGACCTGTGATGAGGCTGCGCCGAGTGGCCAGATGAGGAGCAG 960
 Db 901 GAGTGTGTGCTGACATGTGACCTGTGATGACATGATGATGATGATGATGATGATG 960
 Oy 961 GTGCTTGGCTTCTCAGAGAGGAGTCCAGCTGACAGATTCGCCAGTAGCTGAAGAGC 1020
 Db 961 GTGCTTGGCTTCTCAGAGAGGAGTCCAGCTGACAGATTCGCCAGTAGCTGAAGAGC 1020
 Oy 1021 CACCTGCGCTTGGCCAGCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
 Db 1021 CACCTGCGCTTGGCCAGCAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080

[illegible]

Dd	2152	GTCGCTGCCACAGAGAGTACTGGAGCACTGCACCAGCGGTTCCTGGCTCAGCGCTGGCGTTG	2211
Oy	2197	GGCGACGCCAACCAATGCCACGCTGGCCCTTCTCTGCTTCCTGGCACTTTCCTGEGCGG	2256
Dd	2212	GTCGCACTCACCAACAATGTAATGCTTCCCTGCTTCTGGGCACTTTCCTGGTAGAG	2271
Oy	2257	AGCCAGCGGGGGTGCTAACAACGGTGGCCGCGGCGCCCACTTTCGCAATGCTGGCCCTACTTC	2316
Dd	2272	AGCCAGCGCTGGCGCGCTAACAACCGTBCCCGTGGTCTCACTTCGCGCATGCTAGCTTATTTC	2331
Oy	2317	ATCACCTGGGTCCTCTTGTGGCCCTCCGTGGCAAATGTGAGTGTCTTCAGGCCGCC	2376
Dd	2332	ATCACCTGGGTCCTCTTGTGGCCCTCCGTGGCAAATGTGAGTGTCTTCAGGCCAGCT	2391
Oy	2377	GTCGAGATGGGCGCCCTCTGCTGTGTCTCCGTGGCAATCCGCGCTTCCTACCTGCC	2436
Dd	2392	GTCGAGATGGGCGCTATCTGATCTGTCTCCGTGGCAATCCGCGCTTCCTACCTGCC	2451
Oy	2437	AGGTGTACTCTGCTATCGGACAGCCAGGCGTCAACACCCCGAGTTCTTCTGGAGAGG	2496
Dd	2452	AAGTGTACTGTGCTCTTCTTGGCTGGCAAGCTCAACACCCAGAGTCTTCTGGAGAGG	2511
Oy	2497	GGCCCTGGGATGCCCAAGGCCAGA	2522
Dd	2512	AATGCCAAGAAAGCAGCAGATGAGAA	2537
 RESULT 14 AAD34458 ID AAD34458 standard; cDNA; 1113 BP. XX AC AAD34458; XX DT 16-JUL-2002 (first entry) XX DE Human GPCR cDNA #1.			
XX	Human; G protein-coupled receptor; GPCR; central nervous system disease;		
KW	Cancer; anorexia; bullimia; asthma; benign prostatic hypertrophy; ulcer;		
KW	allergy; acute heart failure; myocardial infarction; urinary retention;		
KW	Parkinson's disease; Alzheimer's disease; hypotension; angina pectoris;		
KW	hypertension; osteoporosis; inflammation; infection; dementia; diabetes		
KM	leukemia; lymphoma; gene therapy; gene; ss.		
OS	Homo sapiens.		
FH	Key	Location/Qualifiers	
FT	CDS	1..1113	
FT		/*tag= a	
FT		/product= "Human GPCR protein #1"	
PN	WO200224885-A2.		
PD	28-MAR-2002.		
PE	21-SEP-2001; 2001MO-EPI0929.		
PR	25-SEP-2000; 2000US-234982P.		
PA	(FARB) BAYER AG.		
PI	Kossida S;		
DR	WPI: 2002-339949/37.		
DR	P-PSDB; AAE21817.		
PT	New G-protein coupled receptor polynucleotide and its encoded protein,		
PT	useful for identifying modulators of GPCR activity, and in gene therapy		
PT	for treating bacterial infection, cancer, acute heart failure or		
PT	inflammation -		
PS	Claim 19; Fig 1; 101pp; English.		

CC The invention relates to an isolated polypeptide, comprising a variant
CC form of mouse or human SACL polypeptide. The variant form is associated
CC with altered preference for carbohydrates, other sweeteners or ethanol.
CC The polypeptide and its associated DNA sequence can be produced by
CC recombinant techniques and is useful for preventing obesity, diabetes or

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 23, 2003, 12:30:29 ; Search time 3569.38 Seconds

(without alignments)
11611.043 Million cell updates/sec

Title: US-09-927-315-14

Perfect score: 2559

Sequence: 1 atgtctggccctgtctgtctct.....atcaggggaacatgagtga 2559

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

EST:*

1: em_estba:**
2: em_esthum:**
3: em_estnu:**
4: em_estov:**
5: em_estpl:**
6: em_estro:**
7: em_estl:**
8: gb_estl:**
9: gb_est2:**
10: gb_est3:**
11: gb_est4:**
12: gb_est5:**
13: gb_est6:**
14: gb_est7:**
15: em_estfun:**
16: em_estom:**
17: gb_gss:**
18: em_gss_hum:**
19: em_gss_inv:**
20: em_gss_pln:**
21: em_gss_vit:**
22: em_gss_fun:**
23: em_gss_mam:**
24: em_gss_mus:**
25: em_gss_other:**
26: em_gss_pro:**
27: em_gss_rnd:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	367.8	14.4	750 14	BQ178590 UI-M-EVO-
2	348.4	13.6	558 12	BE754542
3	219.2	8.6	323 10	AW417577 34413 MAR
4	108	4.2	973 17	CNS02BDU
5	107.4	4.2	938 13	BI462667 603202107
6	106.6	4.2	612 13	BM439858 pgr1n.pk0

7	106	4.1	495	14	BM706305
8	104.8	4.1	298	9	AI562167
9	104.6	4.1	709	12	BG182090
10	102.2	4.0	168	9	AA907022
11	101.8	4.0	1038	17	CNS05068
12	99.2	3.9	464	9	AA937218
13	99.2	3.9	498	9	AM051287
14	98.6	3.9	635	10	AM771483
15	97.2	3.8	738	12	BG191110
16	96.8	3.8	694	12	BG194102
17	86	3.4	830	17	AZ750699
18	78.6	3.1	1041	17	CNS012X1
19	75.6	3.0	829	17	CNS04565
20	74	2.9	718	17	AZ972907
21	73	2.9	902	17	CNS04K5T
22	71.4	2.8	573	9	AI657897
23	69.2	2.7	525	17	AZ483105
24	69	2.7	1470	12	BE910284
25	68.4	2.7	602	9	AI434785
26	67	2.6	562	9	AI742401
27	66.6	2.6	925	17	CNS0091P
28	66.6	2.6	1014	17	CNS033XV
29	65	2.5	487	13	BM259415
30	64.6	2.5	479	13	BM259601
31	63	2.5	935	17	CNS006XK
32	62.6	2.4	571	10	AM453329
33	62.6	2.4	1244	12	BG846745
34	62.4	2.4	1085	17	AG152796
35	62.2	2.4	1034	11	AY111284
36	62	2.4	820	10	BE558912
37	61.2	2.4	545	12	BG414601
38	61.2	2.4	656	13	BI874236
39	61.2	2.4	700	9	AI507686
40	61.2	2.4	810	12	BF628952
41	61.2	2.4	925	17	CNS0091P
42	61	2.4	635	14	BQ768749
43	60.8	2.4	888	17	CNS03200
44	60.6	2.4	454	17	AO466269
45	60.4	2.4	657	14	BQ245583

ALIGNMENTS

RESULT 1
BQ178590/c 750 bp mRNA linear EST 30-APR-2002
UI-M-EVO-bws-h-23-0-UI-r1 NIH.BMAP_EVO Mus musculus CDNA clone
IMAGE:5701318 5', mRNA sequence.
BQ178590
BQ178590.1 GI:20354082
EST.
house mouse.
Mus musculus
ORGANISM
SOURCE
REFERENCE
1 (bases 1 to 750)
NIH-MGC <http://mgc.ncl.nih.gov/>.
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILM at: <http://image.llnl.gov>
(BMAP)
This clone was contributed by the Brain Molecular Anatomy Project
Seq primer: pYX-5.
Location/Qualifiers

FEATURES

QY 1457 GCTGGACACGCTGTGACACACGAGAGCCGTCCTCCGGTGTCTCGGCGAGTCCAGAGAG 1516
 DB 181 TGTGGACACACCCCGAGAGAGAGCCGTCCTCCAGTGTCTACAGGCGAGTCCAGAGAG 240
 QY 1517 GCCAGGTGCGCGGCTCAAGGGGCTTCCACTCTGTCTGTACAGTGTGTGACTCGAG 1576
 DB 241 GCGAGGTGCGCGGCTCAAGGGGCTTCCACTCTGTCTGTACAGTGTGTGACTCGAG 300
 QY 1577 GGGGAGCTACGGGAGAAACGAGAGAGATGCGGCTGTCCTTTGGCCAGAGAGT 1636
 DB 301 CAGGAGCTACGAG 360
 QY 1637 GGTCCCGGAG 1696
 DB 361 GGTCCCGGAG 420
 QY 1697 AGCCGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTG 1756
 DB 421 AGCCGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTG 480
 QY 1757 CTTTGGGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTG 1816
 DB 481 CCTTGGGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTG 540
 QY 1817 TGGCCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTG 1834
 DB 541 GGGCCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTG 558

RESULT 3
 AM417577 323 bp mRNA linear EST 09-JUL-2000
 LOCUS 54413 MARC 1P1G Sus scrofa cDNA 5', mRNA sequence.
 DEFINITION AM417577
 VERSION AM417577.1 GI:6945459
 KEYWORDS EST.
 SOURCE pig.
 ORGANISM Sus scrofa
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.

REFERENCE
 AUTHORS 1 (bases 1 to 323)
 Fahnenkrug, S.C., Fickling, B.A., Rohrer, G.A., Smith, T.P.L., Casas, E.,
 Stone, R.T., Heaton, M.P., Grosse, W.M., Bennett, G.A., Laegreid, W.W.,
 and Keeler, J.W.
 Design and use of two pooled tissue normalized cDNA libraries for
 EST discovery in swine
 Unpublished (2000)
 CONTACT: Smith TPL
 USDA, ARS, US Meat Animal Research Center
 PO Box 166, Clay Center, NE 68933-0166, USA
 Tel: 402 762 4366
 Fax: 402 762 4390
 Email: smithemall.marc.usda.gov
 Single pass sequencing. Bases called and trimmed with phred
 v0.980904.e. Vector identified by cross_match with the -minscore 20
 and -mismatch 12 options.
 PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTTCCAGTACGAGAG
 Plate: 20 row: E column: 14
 Seq primer: ATTTAGGTGACATAATAG.
 Location/Qualifiers
 1..323
 /organism="Sus scrofa"
 /db_xref="taxon:9823"
 /clone_lib="MARC 1P1G"
 /tissue_type="pooled"
 /lab_host="DH10B"
 /note="Vector: pCMV SPORT6; Site 1: NotI; Site 2: SalI;
 Library made from pooled tissue from day 11, 13, 15, 20,
 and 30 embryos."

BASE COUNT 43 a 120 c 85 g 72 t 3 others
 ORIGIN

Query Match 8.6%; Score 219.2; DB 10; Length 323;
 Best Local Similarity 82.3%; Pred. No. 2.1e-29;
 Matches 251; Conservative 0; Mismatches 54; Indels 0; Gaps 0;

QY 2127 GAGCTGGACATGCTGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2186
 DB 19 GGACTGTGGCTGCTACACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 78
 QY 2187 CTTGGGCTTACGCGACGCGACCAATGACAGCTGTCTGTCTGTCTGTCTGTCTG 2246
 DB 79 CTTGAGCCTGTGTACATGTGTGACCAATGACAGAGAGAGAGAGAGAGAGAGAGAG 138
 QY 2247 CTTGGTGGAG 2306
 DB 139 CTTGTGTGACAG 198
 QY 2307 GGCCTACTTCAATCAGCTGGGCTCTCTTGTGTGCCCCCTCTGAGCAATGTGACAGTGTCT 2366
 DB 199 GGCTACTTCAATCAGCTGGGCTCTCTTGTGTGCCCCCTCTGAGCAATGTGAGGCTG 258
 QY 2367 CAGGCGCCGCTGACAG 2426
 DB 239 CCAGCCTGCGGACAGATGGGTGCGCCCTCTCTGTGTGTGTGTGTGTGTGTGTGT 318
 QY 2427 CCACC 2431
 DB 319 CCACC 323

RESULT 4
 CNS02B0U 973 bp DNA linear GSS 12-MAY-2000
 LOCUS CNS02B0U
 DEFINITION Tetraodon nigroviridis genome survey sequence pUC-ori end of clone
 254011 of library G from Tetraodon nigroviridis, genomic survey
 sequence.
 ACCESSION AL190335.1 GI:7828439
 VERSION AL190335
 KEYWORDS GSS: genome survey sequence.
 SOURCE Tetraodon nigroviridis.
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodon.

REFERENCE
 AUTHORS 1 (bases 1 to 973)
 Roest-Crollius, H., Jallou, O., Dasilva, C., Bouneau, L., Fisher, C.,
 Bernot, A., Fzames, C., Wincker, P., Brothier, P., Queller, F.,
 Saurin, W. and Weissbach, J.
 Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence
 Unpublished
 2 (bases 1 to 973)
 Roest-Crollius, H., Jallou, O., Dasilva, C., Fzames, C., Fisher, C.,
 Bouneau, L., Billaud, A., Queller, F., Saurin, W., Bernot, A. and
 Weissbach, J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis

TITLE
 JOURNAL Unpublished
 REFERENCE 3 (bases 1 to 973)
 AUTHORS
 TITLE Direct Submission
 JOURNAL Submitted (12-APR-2000)
 COMMENT This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/Tetraodon.
 Location/Qualifiers
 1..973
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone_lib="254011"
 /clone_lib="G"

FEATURES
 source
 1..973
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone_lib="254011"
 /clone_lib="G"

VERSION	BK439858.1	GI:18470633
KEYWORDS	EST.	
SOURCE	chicken.	
ORGANISM	Gallus gallus	
REFERENCE	Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae; Gallus	
AUTHORS	1 (bases 1 to 612)	
TITLE	Cogburn, L.A. and Nys, Y.	
JOURNAL	Ests from Normalized Chicken Reproductive Tract cDNA Library-University of Delaware and INRA, Tours-Poultry Unit Project Unpublished (2002)	
COMMENT	Contact: Larry A. Cogburn University of Delaware Townsend Hall, Newark, DE 19717, USA Tel: 302-831-1335 Fax: 302-831-2832 Email: cogburnudel.edu, www.chickest.udel.edu.	
FEATURES	Location/Qualifiers	
Source	1..612	
	/organism="Gallus gallus"	
	/strain="Commercial broiler and layer"	
	/db_xref="taxon:9031"	
	/clone="pgrln.pk001.h19"	
	/clone_lib="Normalized Chicken Reproductive Tract cDNA library (pgrln)"	
	/sex="Male and Female"	
	/tissue_type="Testis, ovary and oviduct"	
	/dev_stage="Various stages; embryonic, post-hatch, immature and sexually-mature"	
	/lab_host="E. coli EMEDH10B"	
	/note="Vector: pCMVSPORT6; Library made from three total RNA pools from each tissue (testis 25%, ovary 25%, and oviduct 50% of final RNA pool); Single pass sequencing from 5'-end"	
BASE COUNT	78 a 241 c 202 g 86 t 5 others	
ORIGIN		
Query Match	4.2%; Score 106.6; DB 13; Length 612;	
Best Local Similarity	57.5%; Pred. No. 2.6e-09;	
Matches 206; Conservative	0; Mismatches 149; Indels 3; Gaps 1;	
OY	ACGGCTCCTGTGGCAATGGCCATGAATAAGCGCGTGAGAGATCAACAACAAAGTCGG 262	
DB	ACGGGTACACACTCTCGCAAAATGAGGCCCTTCCGCGTGAGAGATCAACAACACTCACAGC 316	
OY	ATCTGCTCCCGGGCTGGCGCTGAGGCATGACGACGCTTTGATACGTCGTCGAGCGCTGNG 322	
DB	CGCTCTCTCCCAACGTCACGCTGGGGCTGACGATCCACGACACCTGACAGAGGC--GG 373	
OY	TGGCCATGAAGCCCAAGCCTCATGTCTCGGCCAAGGACGAGCGCGACATCGCCGCT 382	
DB	CCAACCTGACAGGACGCTGGCGCGGCTCGGCCGAGGAGGCGGACGACACTGAGAGTGC 433	
OY	ACTGCACTACACGACGATACAGCCCCGTGTGCTGTGCTGATCGGGGCCCACTGCTAG 442	
DB	TGAGGGCCCCCAGGCGATACAGACCCCGCGCGTGGCGCATCGGCCCCAGACATCC 493	
OY	AGCTTGGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 502	
DB	AGCTTGGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 553	
OY	GTGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 560	
DB	AAGCTTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 611	
RESULT 7	495 bp mRNA linear EST 28-FEB-2002	
LOCUS	BM706305	
DEFINITION	UT-E-DW0-agh-a-22-0-UT_r1 UT-E-DW0 Homo sapiens cDNA clone	
ACCESSION	BM706305	

VERSION	BW706305.1	GI:19019563
KEYWORDS	EST.	
SOURCE	human.	
ORGANISM	Homo sapiens	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
TITLE	1 (bases 1 to 495)	
JOURNAL	Bonaldo,M.B., Lennon,G. and Soares,M.B.	
MEDLINE	Normalization and subtraction: two approaches to facilitate gene	
COMMENT	discovery	
	Genome Res. 6 (9), 791-806 (1996)	
	97044477	
	Contact: Soares, MB	
	Program for Rat Gene Discovery and Mapping	
	University of Iowa	
	451 Eckstein Medical Research Building Iowa City, IA 52242, USA	
	Tel: 319 335 8250	
	Fax: 319 335 9565	
	Email: msoares@blue.weeg.uiowa.edu	
	Tissue Procurement: Dr. Gregg Haegem	
	CDNA library preparation: Dr. M. Bento Soares, University of Iowa	
	CDNA library arrayed by: Dr. M. Bento Soares, University of Iowa	
	DNA Sequencing by: Dr. M. Bento Soares, University of Iowa	
	Clone Distribution: Researchers may obtain clones from Research	
	Genetics (www.resgen.com).	
	Seq primer: M3 Reverse.	
FEATURES	Location/Qualifiers	
source	1..495	

```

/lab host="DH10B (Life Technologies) (T1 phage resistant)
/note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-DW0 is a cDNA library containing the following
tissue(s): lens. The library was constructed according to
Bonaldi, Lennon and Soares, Genome Research, 6:791-806,
1996. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was ligated to an EcoR I adaptor, digested with Not
I, and cloned directionally into pT73-Pac vector. The
oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dr)18 tail. The
sequence tag for this library is CGATTACGA. This library
was created for the program, Gene Discovery in the Visual
System, supported by National Eye Institute (NEI)."
BASE COUNT      109 a      178 c      137 g      71 t
ORIGIN

Query Match      4.1%; Score 106; DB 14; Length 495;
Best Local Similarity 100.0%; Freq. NO. 3.3e-09;
Matches 106; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2454 GCGGAGCGCCAGGCGTCACACACCCCCGAGTTCTCTCTGGAGAGGGGCCCTGGGAGATGCCCA 2513
        |||||||
Db 1 GCGGACGCCAGGCGTCACACACCCCCGAGTTCTCTCTGGAGAGGGGCCCTGGGAGATGCCCA 60

OY 2514 AGGCCAGAAATGACGGGAACACAGGAAATCAGGGGAAACATGATGA 2559
        |||||||
Db 61 AGGCCAGAAATGACGGGAACACAGGAAATCAGGGGAAACATGATGA 106

RESULT 8
A1562167/c
LOCUS      298 bp      mRNA      linear      EST 25-MAR-1999
DEFINITION w73d10.x1 Stratagene mouse heart (4937316) MUS musculus cDNA clone
IMAGE:1260595.3 similar to SM:CSR_HUMAN P41180 EXTRACELLULAR
CALCIUM-SENSING RECEPTOR PRECURSOR ;, mRNA sequence.
ACCESSION  A1562167

```

VERSION	AT562167.1	GI:4513512
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	

REFERENCE	AUTHORS
1 (bases 1 to 298)	Kucaba, T., Martin, J., Beck, C., Wylie, T., Marra, M., Hillier, L.

TITLE
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y.,
B., Swaller, T., Gibbons, M., Page, D., Harvey, N., Schurr, R.,
E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R., and Wilson, R.
The Washn-Nt Mouse est protect 1000

TITLE	JOURNAL	COMMENT
the Mashu-NCI Mouse EST Project 1999	Unpublished (1999)	Contact: Marra M/Mashu-NCI Mouse EST Project 1999

Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel.: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MCI:66347

This clone was previously sequenced on the 5' end only, this new data is from the 3' end

Possible reversed clone: similarity on wrong strand

Possible reversed clone: polyT not found

High quality sequence stop: 286.

FEATURES	Location/Qualifiers
source	1. .298

```

/organism="Mus musculus"
/strain="NIH Swiss"
/db_xref="taxon:10090"
/clone="IMAGE:1260595"
/clone_id="Stratagene mouse heart (#37316)"
/sex="pooled"
/tissue_type="heart"
/dev_stage="13 day embryos"
/lab_host="SOLR (kanamycin resistant)"
/note="Organ: heart; Vector: pBluescript SK-; Site.1:
EcoRI; Site.2: XhoI; Cloned unidirectionally. Primer:
Oligo dT. 93 pooled NIH/Swiss 13 day embryo hearts.
Average insert size: 1.0 kb; Uni-ZAP XR Vector: ~5'
adaptor sequence: 5' GAATTCGCACGAC 3' ~3' adaptor
sequence: 5' CTCGACCTTTTCTTTTTTTTTT 3'"
59 a 101 c 87 g 51 t

```

Query Match	4.1%	Score 104.8	DB 9	Length 298
Best Local Similarity	63.5%	Pred. No. 5e-09		
Matches 160	Conservative	0	Mismatches 92	Indels 0
			Gaps	
QY 487	CCCCAGGCGACGTACGGTGTACGATGAGACGTCTGAGCGCCCGGGAGACCTTCCCTCC	546		
Db 298	CCACAGGTCACGTATGAGGCGACGACGCGTATCTCTAGTGGAGAGCGCAAGTTCCTGTC	239		
QY 547	TTCTTCGCGACCGGCGCCACGACCGGTCGTCACGCTGACGGCGCGCGAGGTCGTGAC	606		
Db 238	TTCTTCGCGCACACATCCCGACGCAATAGTACCAAGCTGAAATCTATAGTGGCGCTGCTCAG	179		
QY 607	GAGTTTCGGCTGGAACATGCGGTGGCGCCGCTTGCGGACGACGACGAGTACGGCCCGCAGGAC	666		
Db 178	AGCTTCGGCGTGGGCTGATCTCGCTCGTTGGGAGCATATGTGTAGTACACGCGGACGCTGGCC	119		
QY 667	CTGAGCATCTTTCGTGGCGCTGCGCCGCGGCGACGCGGACATGTGATCGCGGACAGAGGCGCTG	726		
Db 118	GTACAGCGCGCTGGAGGAGCATGGCGCACTCCACGCGGCGCATGTGGTGCCTTCAAGAGACGTG	59		
QY 727	GTGCGCGCTGCC 738			
Db 58	GTGCGCTGTGCC 47			

RESULT 9				
BG182090				
LOCUS	BG182090	709 bp	mRNA	linear
DEFINITION	RS1954 Atherosys RAGE Library Homo sapiens cdna, mRNA sequence.			

KEYWORDS	EST.
SOURCE	human.

ORGANISM	REFERENCE
Homo sapiens	
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.	
1 (bases 1 to 709)	
Hartshorn T T	Shore B
	Bundlitt S
	Talbot B
	Davitt B

AUTHORS
Harrington, J. J., Sherif, B., Rundlett, S., Jackson, P. D., Perry, R.,
Cain, S., Leventhal, C., Thornton, M., Ramachandran, R., Whittington, J.,
Lerner, L., Costanzo, D., McElligott, K., Boozer, S., Mays, R., Smith

TITLE	'E', Veloso, N., Kilka, A., Hess, J., Cuthren, K., Lo, K., Oltropacher, J., Danzig, J. and Ducar, M.
JOURNAL	Creation of genome-wide protein expression libraries using random
MEDLINE	activation of gene expression
COMMENT	Nat. Biotechnol. 19 (5), 440-445 (2001)
	21227151
	Contact: Scott J. Cain
	Athersys, Inc.

3301 Carnegie Ave, Cleveland, OH 44115, USA
Tel: 216 431 9900
Fax: 216 361 9596
Email: scaine@thersys.com
High quality sequence stop: 485.

FEATURES	Location/Qualifiers
source	1. .709

BASE COUNT	217 a	143 c	139 g	209 t	1 others
ORIGIN					
Query Match	4.1%	Score 104.6;	DB 12;	Length 709;	
Best Local Similarity	52.2%	Pred. No. 6e-09;			
Matches 258;	Conservative 0;	Mismatches 230;	Indels 6;	Gaps 1	

[illegible]

```

Db      414 CTGATTCAGAAATCTGGTTGCAGACTCGATTGTGCATCATCAATACCAACAGATGATCAGTATGA   473
Oy      658 CGGCAGGGCCCTGAGCATCTTCTCGGCCCTTGCCGCCGACAGCGGCATCTGCATCGCCGAC   717
          || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      474 CGATTGGCTCTTAACAACCTTTATTAATTACAGGCTGAAGCANMATAACGTGTGATACCTTTC   533
Oy      718 GAGGCGCTGATGCC   731
          | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db      534 AAGAAGCTTCTTC   547

RESULT 10
LOCUS   AA907022
DEFINITION o192a08.s2 Soares_NFL.T.GPC.S1 Homo sapiens cDNA clone
IMAGE:1505750 3', mRNA sequence.
ACCESSION AA907022
VERSION   AA907022.1 GI:3042482
KEYWORDS EST.
SOURCE   human.
ORGANISM Homo sapiens
           Eumariotia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 168)
AUTHORS NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE    National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
           Tumor Gene Index
           Unpublished (1997)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: ccgaps-femail.nih.gov
           This clone is available royalty-free through LINT ; contact the
           IMAGE Consortium (info@image.llnl.gov) for further information.
           Insert Length: 825 Std Error: 0.00
           Seq primer: -40ml3 fwd. Et from Amersham
           High quality sequence stop: 157.

FEATURES
        source
            location/Qualifiers
                1..168
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:1505750"
                /clone_1id="Soares_NFL.T_GBC_S1"
                /lab_host="DH10B"
                /note="Organ: pooled; Vector: pTZ19-Pac (Pharmacia) with
                    a modified polylinker; Site_1: Not I; Site_2: Eco RI;
                    Equal amounts of plasmid DNA from three normalized
                    libraries (fetal lung NDHL19W, testis NTI, and B-cell
                    NCI-GAP-GCB1) were mixed, and ss circles were made in
                    vitro. Following HAP purification, this DNA was used as
                    tracer in a subtractive hybridization reaction. The driver
                    was PCR-amplified cDNAs from pools of 5,000 clones made
                    from the same 3 libraries. The pools consisted of
                    I.M.A.G.E. clones 297480-302087, 682632-687239,
                    726408-728711, and 729096-731399. Subtraction by Bento
                    Soares and M. Fatima Bonaldo."
                25 a 52 c 49 g 42 t

ORIGIN
Query Match 4.0%; Score 102.2; DB: 9; Length 168;
Best Local Similarity 97.2%; Pred. NO. 1.4e-08;
Matches 104; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1171 TTCTCTGTCTACGAGCGTGTATNAGCTGTGCCCCAGGCCCTGCACACACTTTCAGTGC 1230
       . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
Db 11 TTTTTTGTCTACGAGCGTGTATNAGCTGTGCCCCAGGCCCTGCACACACTTTCAGTGC 70

Oy 1231 AACGCTTAGGCTGCGCCGCGCAGACCCTCGTAAGCCCTTGAGCT 1277
       . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . | . |
Db 71 AACGCTTAGGCTGCGCCGCGCAGACCCTCGTAAGCCCTTGAGCT 117

RESULT 11
LOCUS   CNS0506R/c

```

[illegible]

/note=Organ: pooled; Vector: pT733-Pac (Pharmacia) with a modified polynucleotide; Steel: Not I; Site 2: Eco RI; Equal amounts of plasmid DNA from five normalized libraries were mixed, and six circles were made *in vitro*. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the following libraries and clones: Soares NDBst pool 1:

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2555206"
/clone_lib="Soares NSF_F8_9W_OT_PA_P_S1"

```

RESULT	14
LOCUS	AM771483/c
DEFINITION	635 bp mRNA linear EST 04-MAY-2000
ACCESSION	U09714
VERSION	1
KEYWORDS	CDNA SEQUENCE
SOURCE	Homo sapiens
ORGANISM	human.
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 635)
AUTHORS	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997)
JOURNAL	Contact: Robert Strausberg, Ph.D. Email: cgapds@email.nih.gov
COMMENT	

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3031976"
/clone_1lb="NCI CGAP_Kid11"
/lab_host="DH10b"
/note="Organ: kidney; Vector: p7T73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not 1; Site_2: Eco RI;
Plasmid DNA from the normalized library NCI CGAP Kid3 was
prepared, and ss circles were made in vitro. Following NAB
pufication, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneids 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."
BASE COUNT      188 a      126 g      193 t
ORIGIN

```

Query Match	3.9%	Score 98.6:	DB 10;	Length 635;
Best Local Similarity	51.38;	Pred. No. 7e-08;		
Matches 258;	Conservative	0;	Mismatches 239;	Indels 6;
			Gaps	1;
QY 253 AACAGTCGATCTCTGCTCCCGGGCTGGCCCTGGGCTACGACTCTTTGATACGGTCTG 312				
Db 635 AACATTTCAACACTCTTATCTGGAGTCAAACTGGGGTATGAAATCTATGACACTTTGTACA 576				
QY 313 GAGCCTGTGTGGCCATGAAAGCCAGCCTCATGTTCTTGCC-----CAAGGCAAGGACGC 366				
Db 575 GAAATCACAGAGGGCAATGGACAGCACTGTGAGTTCTTTCTAAATTCMACTGCTCCAGA 516				
QY 367 CGCGCATCGCGCGCTACTGCACATPACACGGAGTACACGCGCCCGGCTGCTGCTCATC 426				
Db 515 GAAACTGTGGAGTTTAACTGTGACTATTTCACGCTTACATGCCAAGAGTTTAAAGGCTGTCTATA 456				
QY 427 GGGGCCCACTGCTGAGAGCTCGCCATGGTACACGGCAAGTTCTTCAGCTTCTCCTCATG 486				
Db 455 GGTTCGTGGGTACTCGAATAAATACATATGGCTGTCTCCAGAGATGTTGAATTTACAGTCTATG 396				
QY 487 CCCAGGTCACTACGGTGTGCTTACATGAGAGTCTGTAAGCCGCCGGGAGACCTTCCCTCC 546				
Db 395 CCACAGGTGGGTATTGAAATCAACTGCAGAAATTCCTGATGTACAAAAATTCGGTTCCTTCA 336				
QY 547 TTCTTCGCACCGTGGCCCGAGCGACGTGTGAGGTGAGCGAGCGCGCGGAGCTGTGTGAG 606				
Db 335 TTTTTCAGGACTGTGCCAGTGACTCTCCATPAAATTAAGCAATGCGTCCACTGTTAG 276				
QY 607 GAGTTCGGCTGAACTGAGTGGCGCCCTGTGGACAGCAGAGTATGAGCGCGAGGCGC 666				
Db 275 AAATCTGGTTGGAACTGATTTGGCATCTAATACACAGATGATGACTATGAGACGATTTGGCT 216				
QY 667 CTGAGCATCTTCTCGGGCCGTGGCCGGGACCGGCACTGTGCATCGCGCACGAGGGCCCTG 726				
Db 215 CTTAACACTTTTATAATTCAGGCTGAAGCAATAATACGTGTCAATAGCCTTCAAGAGGTT 156				
QY 727 GTGGCGGCGTCCCGGTGGCCGATGA 749				
Db 155 CTTCAGGCTTTCTTCTTGAGTAA 133				

GenCore version 5.1.4_p5_4578
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 23, 2003, 12:51:29 ; Search time 89.0525 Seconds

8812.617 Million cell updates/sec

Title: US-09-927-315-14

Sequence: 1 atgctggccctgtctct.....atcagggaacatgagtga 2559

Scoring table: IDENTITY_NUC

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

```
Minimum DB seq length: 0
```

[illegible]

Post-processing:	Minimum Match	0%
	Maximum Match	100%

Listing first 45 summaries

```
Database :      Issued_Patents_NA:*
1: /com2 6/ntodata/1/1na/5a COMB seq.*
```

```

1: /cgn2_6.ptodata/1/1na/5A.COMB.seq:*
2: /cgn2_6.ptodata/1/1na/5B.COMB.seq:*
3: /cgn2_6.ptodata/1/1na/6A.COMB.seq:*
4: /cgn2_6.ptodata/1/1na/6B.COMB.seq:*
5: /cgn2_6.ptodata/1/1na/PCRTS.COMB.seq:*
6: /cgn2_6.ptodata/1/1na/backfiles1.seq:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	284	11.1	2532	4	US-09-361-531-4	Sequence 4, Appl1
2	279.4	10.9	2293	4	US-09-361-531-3	Sequence 3, Appl1
3	218.6	8.5	2010	4	US-09-361-631-8	Sequence 8, Appl1
4	173	6.8	3809	1	US-08-485-588-3	Sequence 3, Appl1
5	173	6.8	3809	1	US-08-484-565-3	Sequence 3, Appl1
6	173	6.8	3809	2	US-08-480-751-3	Sequence 3, Appl1
7	173	6.8	3809	2	US-08-943-986-3	Sequence 3, Appl1
8	173	6.8	3809	3	US-08-353-784-3	Sequence 3, Appl1
9	173	6.8	3809	3	US-08-484-7198-3	Sequence 3, Appl1
10	173	6.8	3809	4	US-08-546-998-2	Sequence 2, Appl1
11	173	6.8	3809	4	US-08-484-159-3	Sequence 3, Appl1
12	173	6.8	4000	2	US-08-687-889A-2	Sequence 2, Appl1
13	173	6.8	5006	1	US-08-485-565-2	Sequence 2, Appl1
14	173	6.8	5006	1	US-08-484-565-2	Sequence 2, Appl1
15	173	6.8	5006	2	US-08-480-751-2	Sequence 2, Appl1
16	173	6.8	5006	2	US-08-943-986-2	Sequence 2, Appl1
17	173	6.8	5006	3	US-08-353-784-2	Sequence 2, Appl1
18	173	6.8	5006	4	US-08-546-7198-2	Sequence 2, Appl1
19	173	6.8	5006	4	US-08-546-998-1	Sequence 1, Appl1
20	173	6.8	5006	4	US-08-484-159-2	Sequence 2, Appl1
21	161.8	6.3	5275	1	US-08-485-588-1	Sequence 1, Appl1
22	161.8	6.3	5275	1	US-08-484-565-1	Sequence 1, Appl1
23	161.8	6.3	5275	2	US-08-480-751-1	Sequence 1, Appl1
24	161.8	6.3	5275	2	US-08-943-986-1	Sequence 1, Appl1
25	161.8	6.3	5275	3	US-08-353-784-1	Sequence 1, Appl1
26	161.8	6.3	5275	3	US-08-484-7198-1	Sequence 1, Appl1
27	161.8	6.3	5275	4	US-08-484-159-1	Sequence 1, Appl1

28	152.8	6.0	2177	4	US-09-134-513-1	Sequence 1, Appl1
29	146.2	5.7	2188	5	PCR-US93-01642-1	Sequence 1, Appl1
30	144.8	5.7	4134	4	US-09-162-021B-1	Sequence 1, Appl1
31	144.8	5.7	4134	4	US-09-687-477-17	Sequence 17, Appl1
32	133.6	5.2	4131	1	US-08-485-586-4	Sequence 4, Appl1
33	133.6	5.2	4131	1	US-08-484-565-4	Sequence 4, Appl1
34	133.6	5.2	4131	2	US-08-480-751-4	Sequence 4, Appl1
35	133.6	5.2	4131	2	US-08-943-986-4	Sequence 4, Appl1
36	133.6	5.2	4131	2	US-08-353-784-4	Sequence 4, Appl1
37	133.6	5.2	4131	3	US-08-484-179B-4	Sequence 4, Appl1
38	133.6	5.2	4131	3	US-08-484-159-4	Sequence 4, Appl1
39	101.2	4.0	3384	2	US-08-687-289A-1	Sequence 1, Appl1
40	95.8	3.7	4095	5	PCR-US91-09422-18	Sequence 18, Appl1
41	93.8	3.7	2736	4	US-08-617-785-1	Sequence 1, Appl1
42	93.8	3.7	3451	4	US-09-641-318-1	Sequence 1, Appl1
43	93.8	3.7	3431	4	US-09-641-318-3	Sequence 3, Appl1
44	87.4	3.4	2436	5	PCR-US91-09422-20	Sequence 20, Appl1
45	77.8	3.0	2619	3	US-08-337-797A-1	Sequence 1, Appl1

ALIGNMENTS

RESULT 1

; Sequence 4, Application US/09361631

GENERAL INFORMATION:

APPLICANT: Adler, Jon Elliott

APPLICANT: The Regents of the University of California

; TITLE OF INVENTION: Involved in Sensory Transduction
FILE NUMBER: 033027-00023000

CURRENT APPLICATION NUMBER: US/09/361, 633;
CURRENT FILING DATE: 2009-07-27

EARLIER APPLICATION NUMBER: US 60/095,466
EARLIER FILING DATE: 1998-07-29

EARLIER APPLICATION NUMBER: US 60/112,744

```

; NUMBER OF SEQ ID NOS: 10
SOFTWARE: PatentIn Ver 3.0

```

; SEQ ID NO 4
; LENGTH: 3533

```

; TYPE: DNA
; ORGANISM: Mus musculus

```

OTHER TAX

; OTHER INFORMATION: nucleotide sequence
ms-00-363-633-A

Query Match	11.18;	Score 284;	DB 4;	Length 2532;
-------------	--------	------------	-------	--------------

Matches 1180; Conservative 0; Mismatches 1185; Indels 65; Gaps 12,

QY 92 TGAAGGGGGACTACGATGCTGGGGGGGGCTGTTCCTCCCTGGGGAAGCCGAGAG-----G 145

Db 95 TGGCTGGGGACTACTCTCTGGGGTGGGCTTTTACCCTTCATGCCAAGTGAAGGTGCT 154

OY 146 CTGGCTTCGACGCCGGACACGCCCCACAGCCCTGTGTGACCAAGTTCCTCTCAACG 205

Db 155 CTCACCTACGCTACTCTGCAGGTGCCCCAAGTGCAATGAGTACACATGMAAGGTGTGGGCT 214

OY 206 GCCGTCTGTGGGCACTGGGCCATGAAATGCGCGTGGAGGAGATCAACAAGTCGGATC 265

Db 215 ACAACCTCATG---CAGGCGCATGTGGATTGCGCTGGGGGAAATCAACAACGTAGCTCTT 271

OY 266 TGTGTCCCGGGGCTCGGCTGGGGCTACGACCTCTTTAGTAGCTGTCGAGACCTGTGGTGG 325

Db 272 TGTCTCCGGGGGTGCTGTGTGGCTACGAGATGGTGGATGTCTGT---ACCTCTCCACA 328

OY 326 CCATGAAAGCCCAAGCTCATGTTCTCTGGCAAGGACGAGGACGCGGACATCGCGGCTACT 385

Db 329 ATATCCAGCCCTGGGCTCTACTCTCTGTCACAGATAGATGACTCC---TGCCCATCTCTCA 385
 QY 386 GCACCTACAGCAGTACACAGCCCGCTGCTGCTGATCGAGGCGCCACATCGACAGC 445
 Db 386 AAGACTACAGCAGTACAGGCGCCCAAGTGCTGTATATGGCCAGACACTGTGAGT 445
 QY 446 TCAGCATGCTACAGCAGAGTCTTCTGACTTCTCTCATAGCCCGAGTACAGTACGCTG 505
 Db 446 CTGCATACAGCAGTACAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 505
 QY 506 CTACATAGAGACTCTGAGCGCCCGGAGACCTTCCCTCTCTCTCTCTCTCTCTCTCTCT 565
 Db 506 CCATCACCAGCAGTCTGACAGAGGCGCTTCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 565
 QY 566 GCGACCGCTGCTGAGTACAGCGCCCGGAGCTGCTGACAGAGTGGCTGGCTGAGTACG 625
 Db 566 GCGCCACCCACCATGATGAGGCGCATGTGTCACATGATGTTCACCTTCCAGTGGAGTGA 625
 QY 626 TGCGCCGCTGAGGAGCAGCAGATACGCGC---GACAGGCGCTGAGCATCTCTCTCTG 682
 Db 626 TCCTGTGCTGTGAGCGATGACATTAATGCGAGAGAACAGCACCTGCTGAGCGCAGC 685
 QY 683 CCTTGGCCGCGGAGCAGGCGCATCTGATCGCGCAGAGGCGCTGGTGGCGCTGCGCCGCTG 742
 Db 686 GTCTGACCAACAGTGGCGACATCTGATCTTCCAGAGAGTCTTCCCTTACAGAAC 745
 QY 743 CCGATGATCTGCGGCTGGGAGAGTGCAG---ACGTCTCTGACAGCAGTGA 790
 Db 746 CCAACAGAGCTGTGAGGCTTACAGAGCAGAGCAGCACTGAGAACATCTCTGAGCAAGCTGC 805
 QY 791 ACCAGAGCAGCTGAGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 850
 Db 806 GCGGAGCTTCTGGCGCTGT 865
 QY 851 TCACATACAGATACAGAGAGGCTCTGTGCGCCAGAGTGGGTGGCGCCAGAGCTGTGC 910
 Db 866 TCCGTGAGAGTCTCGCGTGGAGCTTCAAGGCTTGTGTGTGTGTGTGTGTGTGTGTGTGT 925
 QY 911 TGACCTTGAACCTGTCTATGAGGCTGCGCCGATGAGCCAGATGAGGAGGAGCTGTGCT 970
 Db 926 CCAATGAGACCTGTCTTACACACTCAGAGAGCTGTGCGCAGAGGAGGAGCTTCTCTG 985
 QY 971 TCCTCCAGAGGAGGCTGCCAGCTGACAGATTTCCCGAGTACGTGAAGAGCAGCTGTGC 1030
 Db 986 TCACCATCAGAGGAGTGTCTATCTCTGCTTCAAGCAGTCTCGAGTGGCGCATGACAGC 1045
 QY 1031 TGCGCACCGAGCCGCTTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1090
 Db 1046 -----CAGGATATGCAATGCTTAAAGAGACCGAGCTCGGAGCTATCC 1086
 QY 1091 ACGTGTGGGCGCAGCAGCTGCGCCAGTGTGACTGATCAGCTGC-AGAAAGTGTAGCGCA 1149
 Db 1087 TGTAAACAGAGTGTGAGAGCTGTGATGAATCACTGATGCTCTTCAACAGATTTCTCATG 1146
 QY 1150 GGGCTAAATCAACACAGAGCTTCTGTCTACGAGCTGTGTATAGCTGTGGCCAGGCGC 1209
 Db 1147 CTTCGAGGAGGAGCTGT 1206
 QY 1210 CTGACAAACATCTTTCAGTGTGAAGCGCTCAGAGCTGCGCCGCGGAGAGCCCGTGAAGCC 1269
 Db 1207 CTCCACAGACTCTCTCAGTGTGAATGAGTCCGCTGCACCC---AGCAATGTGTATCCA 1263
 QY 1270 TGGGAGCTCTGTGAGATGTACAACTGACTTCCAGCTGTGGGCGGCTCCGCTGAGG 1329
 Db 1264 TGGGAGCTCTGTGAGAGATGTGAGATGTCAACTTCAAGCTCTGTGGGAGACAGACTCTTC 1323
 QY 1330 TTGCAACAGCAGGAGAGAGTGTGAGATGTGAGATGTGAAGCTGTGAAGCTGTGGAGGCG 1389
 Db 1324 TTGCAACAGCAGGAGAGAGTGTGAGATGTGAGATGTGAAGCTGTGAAGCTGTGGAGGCG 1383
 QY 1390 TCAGTGGCCAGGCTTCCAGAGTGTGGAGGCTTCAAGCGGAGGCTTCA-----GGACAGAG 1443

Db 1384 AGCCAGAACCCCTTCCAAAGCATTCGCTCTACTTCCCGACGAGAGCAGAGCTGACTAC 1443
 QY 1444 CGCCTGAAGATTCGCTGGCAGAGCTGTGACACAGAGAGCCCGTGTCCGCTGTGCGG 1503
 Db 1444 ATTGAGATGTGTCTGTGTACACCCCAACAGAGGCTCCCATATGATGTGTCTAG 1503
 QY 1504 CAGTGCAGAGAGGCGCAGGTGTGCGCGGAGTGAAGGGTTCACATCTCTCTCTCTCTCTCT 1563
 Db 1504 AGTTGACAGCTGTGGGCAATGAAAAAACCATAGGCTCCACACCATCTCTCTCTCTCTCT 1563
 QY 1564 GTGAGCTGCAAGGCGGAGAGTACCGGCAAAACCCAGAACACATGCGCTGACTTGTGT 1623
 Db 1564 GTGAGCTGCGCGCGGAGACCTTCAACCGATCATGATGATGATTTAACTGTGTGTCTC 1623
 QY 1624 GCGCAGATGATG---GTGTCCCGCGAGCAGAGCAGAGCTGTCTCCGCGAGTGTCTG 1680
 Db 1624 TGCCCGGCTTCCATGT 1683
 QY 1681 TTCTGTGATGAGGCGAGCGGCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 1740
 Db 1684 TTCTGTGATGAGCAGAGAGT 1743
 QY 1741 GCGCTGT 1800
 Db 1744 ATGATGAGCTGTGAGATGT 1803
 QY 1801 GCTCTGGGAGGCGCCCTGTGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1860
 Db 1804 TGCGCGGAGGCGCCCATATGT 1863
 QY 1861 GTCTCTGT 1920
 Db 1864 GTCCCGCTGT 1923
 QY 1921 CACCTCCCGT 1980
 Db 1924 ACCGT 1983
 QY 1981 GAGTCAAGATGTCTGT 2040
 Db 1984 GTCTTCAAGATGT 2043
 QY 2041 GCTGT 2100
 Db 2044 CCTTACGTCTTGT 2103
 QY 2101 GTGCGCTTCCCGCGGAG---GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2157
 Db 2104 CTGCGCACACATCAACCCCATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2163
 QY 2158 GTGCACTGCGCAGCAGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 2217
 Db 2164 CTCTGT 2223
 QY 2218 CTGAGCTTGT 2277
 Db 2224 CTGT 2283
 QY 2278 GGTGCGGCTGT 2337
 Db 2284 GAGGCAAGT 2343
 QY 2338 CCGCTCTGT 2397
 Db 2344 AGT 2403
 QY 2398 CTGT 2457
 Db 2404 CTCACTTGT 2463
 QY 2458 CAGCCAGGAGCTCAACCCCGAGTCTTC 2487
 Db 2464 TACCGAGGCGCAACCTTCACTTATTTTC 2493


```

QY 1742 GCGTTGCTGGGCTGCTTTGGGCTGTTGTTACCATCGGAGACCCACTGGTTCAGG 1801
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1757 TCAGTACACTGGCCATTCTTTTCATCTTCTGGAGACATTTCCAGACACCCATGGTGGCT 1816
QY 1802 CCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1861
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1817 CGGGCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1876
QY 1862 TCTCTCTGTTCCCTGGCCAGCCCGCCAGCCCTGCGGATGCTGGCCAGGACCCCTTGTGCC 1921
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1877 TGCCCCGTGTATGTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 1936
QY 1922 ACCTCCGCTGACGGGGCTGCTGTGACACACTTTCCTGACAGGGGGGGGGGGGGGGGG 1981
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1937 CCGTGTGCTTCTCATCTGCTATCTGATCTGACCGGTCGCTGCTCCATGAGATCGTGTG 1996
QY 1982 AGTCAGAACTGCTCTGTAGCTGGGGGAGACCGGGCTGAGTGGCTGCTGGGGGGGGGG 2041
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1997 TCTTCAAGATGGCCAGACGGCTGCGCAAGTGCCTACAGTTTTTGGATGGGTTACACAGGG 2056
QY 2042 CCTGGCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2101
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2057 CCTATGTCTTCTGCTGCTGCTTATCAAGGGCATCAAGTGGCTGCTGCTGCTGCTGCTG 2116
QY 2102 TGCCCTTCCCGCCGAGAGTGTGACGGACTGGACATGCTGCCACAGGAGCGCTGGTGC 2161
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2117 TGCGCCACCATCATCAACCCCATTTGGCGGAGCCGACCCGATGACCCCAACATCATGATCC 2176
QY 2162 ACCTCCGACACAGCGCTCCGGGCTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2218
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2177 TCTGTGCTGACCTTACTACCTGACGAGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2236
QY 2219 TGCCCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2278
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2237 TGTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2296
QY 2279 GTGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2338
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2297 AAGCCAACTTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2356
QY 2339 CCGTCCGCGCAATGTGAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2398
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2357 CTTCTATGTCTGTGACAGCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2416
QY 2399 TCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2458
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2417 TCACTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2476
QY 2459 AGCCAGGGCTTAACACCCCGAGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2518
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2477 ACCCGGAGCGGACACCTGACCTGCTATTTCAATAGCATGATCAGGGGCTACACATGAGA 2536
QY 2519 AGA 2521
      |||
Db 2537 AGA 2539

```

```

; EARLIER FILING DATE: 1998-07-28
; EARLIER APPLICATION NUMBER: US 60/112,747
; EARLIER FILING DATE: 1998-12-17
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 8
; LENGTH: 2010
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: human G-protein coupled receptor (GPCR) B4
; OTHER INFORMATION: nucleotide sequence
US-09-361-631-8

Query Match      8.5%; Score 218.6; DB 4; Length 2010;
Best Local Similarity 46.8%; Pred. No. 5,9e-32;
Matches 941; Conservative 0; Mismatches 1014; Indels 54; Gaps 6;

QY 494 TCACCTACGGTGTAGACATGAGCTGCTGAGCGCCCGGAGACCTTCCCTCTTCTTC 553
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2 TCACCTACAGCGCCATCAGCGATGCTGCGAGCAAGTGTGCTGCTCCGCTTGTGCTGC 61
QY 554 GCACCTGTGCGGAGCGGCTGTGAGCTGAGCGCGCGCGGAGCTGCTGAGAGGTTGC 613
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 62 GTACACACACCGAGCGCCGAGCACCATGCTGAGGCTATGTGTACATGTGTGCTGCTTCC 121
QY 614 GCTGGAACGTGGGTGCGCCCTGCGGAGCGAGCAAGTACGCGCGGAGGCGCTGAGCA 673
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 122 GCTGGAACGTGATCATTTGTGTGTGTGAGCAGCAGACCTTATGCGCGGAGCAATGGCCAG 181
QY 674 TCTTCTGCGGCTGCGCGGAGCGGAGCATCTGATCTGCGGAGGAGGCTGTGCGGC 733
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 182 TGCTTGGCGAGCGGCTGCGCGGCGCGGAGCATCTGATCTGCTTCAAGAGACGCTGCCA 241
QY 734 TGCCCGGCGGATGATCTGCGGCTGGGGAAG-----TCAGGAGCTGCTGC 781
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 242 CACTGCAAGCCCAACCAACATGACGTGAGAGAGCGCCAGCGCTGTGATGCTTGTGG 301
QY 782 ACCAGGTGAACAGAGAGCGGTGAGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 841
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 302 ACAAGGTGAGAGAGAGAGAGCGGCGGTCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 361
QY 842 AGCGCTTCTTCACTACATGACATGACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 901
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 ACACTTCTTCAAGAGAGTGTGCGGAGAACTTCAAGAGGCGCGCTGTGATGCTGCCCG 421
QY 902 AGCGCTGCTGACCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 961
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 AGTCTGCGCATGAGACCGGCTCTGCAACACCTGACAGAGCTGTGGCCTGTGGGCACT 481
QY 962 TGTGTGCTTCTTCAAGAGGAGGCTGCCAGCTGACAGAGTTCGCCAGTACGTGAGAGCG 1021
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 482 TCTGTGCGCATGACATCAGAGCGTGCCTATCCGGGCTTCAAGAGTTCGCCAGTGGG 541
QY 1022 ACCTGGGCTGCGCACCGACCGCGCTTGTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTG 1081
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 542 GC-----CCAGAGCTGGGGCGCCGACCCCTGACAGAGACCAAGCTATATAC 591
QY 1082 TGGAGAGAGAGTGTGGGCGACGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1141
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 592 TGCAACAGAGAGTGTGAGCAACTGCTGACAGCGCACTTGTCTTCAACACATTTCTAGG 651
QY 1142 TGAGCGAGAGGCTAATACCAACCAAGAGCTTCTGTCTACAGCAGCTGTGTATAGCTGG 1201
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 652 CTCTCTGGGAGCGTGTGCTTACAG-----CGTGTACTGCGGCTGATCTGTGG 703
QY 1202 CCGAGCGCTGTGCAACACTTTCATGTGCAACGCTTACGCTGCTGCTGCTGCTGCTGCTG 1261
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 704 CCGATGCTGTGCAACAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 760
QY 1262 TGAAGCCCTGCGACCTGCTGAGAGAACTGTACAACTGACCTTTCACAGTGGCGGCGCTGC 1321
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 761 TCTACCTCTGCGACCTGCTTGTAGAGATCTGTGAGAGTGTCAACTTCTCTGTGACCAAC 820

```

```

RESULT 3
US-09-361-631-8
; Sequence 8, Application US/09361631
; Patent No. 6383778
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Adler, Jon Elliot
; APPLICANT: Lindemeier, Juergen
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
; FILE REFERENCE: 02307E-088720US
; CURRENT APPLICATION NUMBER: US/09/361,631
; CURRENT FILING DATE: 1999-07-27
; EARLIER APPLICATION NUMBER: US 60/095,464

```


TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 373..3606
 OTHER INFORMATION:
 US-08-485-588-3

Query Match 6.8%; Score 173; DB 1; Length 3809;
 Best Local Similarity 51.1%; Pred. No. 1.5e-23;
 Matches 497; Conservative 0; Mismatches 455; Indels 21; Gaps 3;

QY 79 CAGCACTTAGATGAAAGGGGACTAGCTGCTGGGGGCTGTCCCTGGGAGGCC 138
 DB 442 CAGGAGGCCCAAGAGGGGGACATATCCTTGGGGGCTCTTCCATTCATTTTGA 501
 QY 139 GAGAGGCTTGGGCTCCGAGCCGAGCCAGCCAGCCAGCCAGCCAGCCAGCCAG 195
 DB 502 GTAGACAGCTAAAGATCAAGATCTCAAAATCAAGCCGAGAGTCTGTGGAATGTATGAGTAT 561
 QY 196 TCCCTAAACGGCTGCTGTGGGCACTGGCCATGAAATGCGCTGAGAGATCAACAC 255
 DB 562 AATTCGGTGGTTCGCTGTGTTACAGGCTATGATTTTGCATAGAGAGATTAACAGC 621
 QY 256 AAGTCGATCTGCTGCGGGGCTGCGCTGCGCTAGCAGCTCTTGTATGCTGCTGGAG 315
 DB 622 AGCCAGCCCTCTTCCCACTTACGCTGGGATACAGATATTTGACACTTGCACAC 681
 QY 316 CCTGTGGGCGCATGAAGCCGAGCCATGTTCTGCGC-----CAGGCGAGGAGCCGC 369
 DB 682 GTTCTAAGGCTTGGAGCCAGCCAGTGTGTTGCTCAAAACAAATGATTTCTTGG 741
 QY 370 GACATCGCCGCTTCACTCACTACAGCAGATACAGCCGCTGCTGCTGCTGCTGCTG 429
 DB 742 AACCTGATGATGCTCAACTGCTCAGAGCAGATTCCTCAGATGCTGCTGCTGGA 801
 QY 430 CCCCACGCTCAGAGCTGCGCATGCTACCGGCAAGTTCTTCACTTCTTCTCATGCC 489
 DB 802 GCAATGCTCTCAGGCTGCTCAGGCGAGTGGCAATCTGCTGGGCTCTTCTCATATCC 861
 QY 490 CAGTCTCAGCTACGCTGCTAGCATGAGCTGCTGAGCGCCGCGGAGACCTTCCCTCTC 549
 DB 862 CAGTCTCAGTATGCTGCTCTCAGAGACCTCTCAGCAACAAATGATTCATGCTTTC 921
 QY 550 TTCCGCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 609
 DB 922 CTCGGAACCAATCCCAATGATGAGCAGCAGGCTGCTGCTGCTGCTGCTGCTGCTG 981
 QY 610 TTGGGCTGGAAGCTGGGCGCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 669
 DB 982 TTCCGCTGGAAGCTGGGCGCAATGCTGATGATGATGATGATGATGATGATGATGAT 1041
 QY 670 AGCATCTTCTGCGCCCTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 729
 DB 1042 GAGAAATTCGAGAGAGAGAGAGAGAGAGAGATCTGATGATGATGATGATGATGATG 1101
 QY 730 CCGCTGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCTG 789
 DB 1102 TCCAGTACTCTGATGA-----GGAAGAGATTCAGCATGATGATGATGATGAT 1149
 QY 790 AACCAAGAGAGAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 849
 DB 1150 CAAATTCGAGAGAGAGAGAGAGAGAGATCTGATGATGATGATGATGATGATGATG 1209
 QY 850 TTCATCTAGCATCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 909
 DB 1210 ATCAAGAGATTTGCTCCGCGCAATATACGCGCAAGATCTGCTGCGCAGGAGGCTGG 1269
 QY 910 CTGACCTCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 969
 DB 1270 GCCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1329
 QY 970 TTCTCTCAGAGAGGCTGCGCAGCTGACGAGTTCCTCCAGTACGTAAGAGCAGCAGCTG 1029

DB 1330 TTCCTGGAAGCTGGGCGAGATCCAGGCTTCCGGGAATTCGAGAAGGTCCATCC 1389
 QY 1030 CTGGCCACCGCACC 1042
 DB 1390 AGGAGTCTGTCC 1402

RESULT 5 US-08-484-565-3

Sequence 3, Application US/08484565
 Patent No. 5763569
 GENERAL INFORMATION:
 APPLICANT: Edward M. Brown
 APPLICANT: Steven C. Hebert
 APPLICANT: James E. Garrett, Jr.
 TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 MOLECULES
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Lyon & Lyon
 STREET: First Interstate World Center
 STREET: Suite 4700
 CITY: Los Angeles
 STATE: California
 COUNTRY: USA
 ZIP: 90071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: FASTSEQ
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/484,565
 FILING DATE: 7 June, 1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 PRIOR APPLICATION DATA: including application
 PRIOR APPLICATION DATA: described below: 9
 APPLICATION NUMBER: 08/353,784
 FILING DATE: 9 December, 1994
 APPLICATION NUMBER: PCT/US/94/12117
 FILING DATE: 21 October, 1994
 APPLICATION NUMBER: U.S. 08/292,827
 FILING DATE: 23 August, 1994
 APPLICATION NUMBER: U.S. 08/141,248
 FILING DATE: 22 October, 1993
 APPLICATION NUMBER: U.S. 08/009,389
 FILING DATE: 23 February, 1993
 APPLICATION NUMBER: U.S. 08/017,127
 FILING DATE: 12 February, 1993
 APPLICATION NUMBER: U.S. 07/934,161
 FILING DATE: 21 August, 1992
 APPLICATION NUMBER: U.S. 07/834,044
 FILING DATE: 11 February, 1992
 APPLICATION NUMBER: U.S. 07/749,451
 FILING DATE: 23 August, 1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Hebert, Sheldon O.
 REGISTRATION NUMBER: 38,179
 REFERENCE/DOCKET NUMBER: 213/006
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3809 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA to mRNA

FEATURE:
NAME/KEY: CDS
LOCATION: 373..3606
OTHER INFORMATION:
US-08-480-751-3

Query Match 6.88; Score 173; DB 2; Length 3809;
Best Local Similarity 51.18; Pred. No. 1.5e-23;
Matches 497; Conservative 0; Mismatches 455; Indels 21; Gaps 3;

79 CAGCAACTTGAGTGAAGGGAGTACGTGTGGGGGGGCGTCCCGGGGGAGGCC 138
DB 442 CAGCGACCCCAAGAGAGGGGACATTAATCTTGGGGGCTCTTCTATTCATTGGA 501
QY 139 GAGGAGCT--GGCTCCGACGCGGACAGCGCCAGACCCCTGTGTACAGGTTTC 195
DB 502 GTACGACTTAAGATCAAGATCTCAATCAAGCCGAGAGTCTGTGAATGATCAGGAT 561
QY 196 TCCCAAGAGGCTGTCTGTGGGCTGCGCATGAAAATGGCCGTGTGAGAGATCAAC 255
DB 562 AATTCGCTGGGTTCCTGTGTACAGGCTATGATTTGGCATAGAGAGATTAACAGC 621
QY 256 AAGTCGATCGTGGCGGGGCTGGGCTGAGAGACCTTGTGATGAGTGTGGAG 315
DB 622 AGCCCAAGCTTCTTCCCACTTACGCTGGGATACAGATATTGACACTTGCACACC 681
QY 316 CCTGTGTGGCCATGAAGCCAGCCCTCATGTCTCTGCGC-----CAAGGACGAGCCGC 369
DB 682 GTTCTTAAGGCTTGAAGGACACCTGATTTTGTGTCAAAACAAATGATTTTG 741
QY 370 GACTTGGCGGCTTACTGCACTACAGCAGTACAGCCCGCTGTGTGTGTGTGTG 429
DB 742 AACCTTATGATGTCTCAACTGTCTAGAGACATTCCTCTACGATTTGTGTGTGGA 801
QY 430 CCCCACGTCGAGAGTCCGATGTCACCGGCAAGTTCTACGCTCTTCCATGCC 489
DB 802 GCACCTGGCTAGGCTCTCCAGCGCAGTGGCAAAATCTGTGGGGCTCTTCTACATGCC 861
QY 490 CAGGTCACTAGGCTGTAGCATGAGCTGTGAGCCCGGAGACCTTCCCTCTTC 549
DB 862 CAGGTCACTATGCTCTCTCCAGCAGACTCTCTCAGCAAAAGATTCATCACTCTTTC 921
QY 550 TTCCGACCGGCTGAGCAGACCGCTGTGAGCCCGGAGAGCTGTCTCTCAGAG 609
DB 922 CTCGGAACCACTCCCAATGATGAGCAGACCGCCTCATGCGACATCATGAGATAT 981
QY 610 TTGCGCTGAACTGGTGGCGCCCTGGGCGAGCAGCAGAGTACGCGCGGAGCGCTG 669
DB 982 TTCCGCTGAACTGGTGGCGCAATGCGACGCTGATGACGATATGGCGCGCGGAGAT 1041
QY 670 AGCATCTTTCGCGCCCTGGCGCGGCGAGCGGCAATGCAATCGGCGAGAGGCTGTG 729
DB 1042 GAGAAATTCGAGAGGAAGTGAAGGAAGATATGCAATGCACTTCAAGTGAATCATC 1101
QY 730 CCGTGGCGGCTGCGCATGAGCTGGGAGAGGTGAGAGCGTCTGCAACAGTG 789
DB 1102 TCCGAGTACCTGTATGA-----GGAAGATTCAGCATGTGTGATGAGAGTAT 1149
QY 790 AACGAGAGCGGCTGCTGTGCTGTGCTGCTCCGCGACGCGCGCCAGCCCTC 849
DB 1150 CAAATTCAGGCGCAAGATCATGTGTGTTTCTCCAGTGGCCCAAGTCTTGAAGCCCTC 1209
QY 850 TTCAACTAGACATCAGCAGCAGGCTCTCCCAAGGTGTGGTGGGCGAGAGCGCTG 909
DB 1210 ATCAAGAGATGTTCGGGCGCAATATCAGCGGCAAGATGTGGGCGAGAGCGCTG 1269
QY 910 CTGACCTCTAAGCTGTGATGAGGCTCCCGGCGATGGCCCAAGATGGCGACGCTTGGC 969
DB 1270 GCCAGCTCTCCCTGATGCGCATGCTCACTACTTCCACGTGTGTGGCGGCGCATTTGA 1329
QY 970 TTGCTCAGAGGGGTGGCGAGCTGACAGAGTTCCCGCAGTACGAGAGAGCAGCAGCTGGCC 1029
DB 1330 TTGCTCTGAAGGCTGGCGAGATCCAGGCTTCCGGGAATTCCTGAAGAGGCTCATCC 1389

QY 1030 CTGGCCACCGACC 1042
DB 1390 AGGAGCTGTGCC 1402

RESULT 7
US-08-943-986-3
Sequence 3, Application US/08943986
Patent No. 5962314
GENERAL INFORMATION:
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
City: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/943, 986
FILING DATE: 03-Oct-1997
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/484, 565
FILING DATE: 7-June-1995
APPLICATION NUMBER: 08/353, 784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292, 827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141, 248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009, 389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017, 127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934, 161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834, 044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749, 451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Hebert, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 3809 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS


```

Db 622 AGCCAGCCCTTCTCCAACTTGAGCTGGATACAGATTTTGAACCTTGCACAC 681
Qy 316 CCTGTGTGCGCAGTAAAGCCAGCTCATGTTCTCTGGC-----CAAGGACAGGACGCC 369
Db 682 GTTCTTAAGGCGCTTGAGAGCACCCTGAGTTGTGTCTCAAAACAATAATGATCTTTG 741
Qy 370 GACATCGCGCGCTACTGTGAATACAGCAGTACAGCCCGCTGTGCTGTGCTATCGGG 429
Db 742 AACCTTGATGATCTTGAACTGTCAAGACATTTCCCTTAGATGCTGTGTGGGA 801
Qy 430 CCCCAGCTGTGAGAGCTGCGCATGTCACCGGAGTCTTACAGTCTTCTCTCAATGCC 489
Db 802 GCAACTGAGCTGAGGCGTCTCCACGCGCATGCGCAATTCGTCGCGCTTTTACATTTCC 861
Qy 480 CAGGTACAGTACAGGCTGTAGCATGAGAGCTGTGAGCGCCCGGAGACCTTCCCTCTTC 549
Db 862 CAGGTACAGTATGCTCTCTCCAGCAGACTCTCTCAAGCAACAAGAAATCAATTCAGTCTTTC 921
Qy 550 TTCCGACACGTCGCGCAGCAGCAGCTGTGACGTGACGCGCGCGGAGCTGTGAGAG 609
Db 922 CTCGAGACCATCCCAATGATGAGCAGCAGCCACTGCGCAGACATCATGATGAT 981
Qy 610 TTCCGCTGGAAGTGTGTCGCGCCCTGCGCAGCAGCAGTACGCGCGCAGGCGCTG 669
Db 982 TTCCGCTGGAAGTGTGTCGCGCCCTGCGCAGCAGCAGTACGCGCGCAGGCGCTG 1041
Qy 670 AGCATCTTCTGTGCGCGCTGCGCGCGCAGCAGCAGTACGCGCGCAGGCGCTG 729
Db 1042 GAGAAATTTCCAGAGAGAGCTGAGAGAGAGATATCTGATGATGATGATGATGAT 1101
Qy 730 CCGTGTGCGCGCTGCGCAGTATGCGCGCGTGGGAGAGTACGAGTCTGTGACAGG 789
Db 1102 TCCAGTATCTGTATGA-----GAGAGATTCAGCATGTGTGAGAGTGT 1149
Qy 790 AACGAGACAGCTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 849
Db 1150 CAAATTCAGCGCGCAAGATATGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1209
Qy 850 TTCAACTTACATGATGAGAGAGCTGTGCGCCCAAGTGTGTGTGTGTGTGTGTGTGT 909
Db 1210 ATCAAGGAGATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1269
Qy 910 CTGACCTGTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 969
Db 1270 GCGAGCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1329
Qy 970 TTCTTCCAGAGGCTGCGCAGCTGTGAGAGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1029
Db 1330 TTCGCTGAGAGGCTGCGCAGATTCGCGGAAATTCGTAAGAGAGTGTGTGTGT 1389
Qy 1030 CTGCGCACCGAC 1042
Db 1390 AGGAGTGTGTGT 1402

```

RESULT 12

```

US-08-687-289A-2
Sequence 2, Application US/08687289A

```

```

GENERAL INFORMATION:
PATENT NO. 5981195
APPLICANT: Fuller, Forrest H.
INVENTOR: Krapcho, Karen J.
TITLE OF INVENTION: CHIMERIC RECEPTORS AND METHODS FOR
IDENTIFYING COMPOUNDS ACTIVE AT
METABOTROPIC GLUTAMATE RECEPTORS AND
THE USE OF SUCH COMPOUNDS IN THE
TREATMENT OF NEUROLOGICAL DISORDERS
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon

```

```

STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
MEDIUM TYPE: storage
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq for Windows 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/687, 289A
FILING DATE: July 25, 1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/001,526
FILING DATE: July 26, 1995
ATTORNEY/AGENT INFORMATION:
NAME: Walburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 220/004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 4000 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-687-289A-2

```

```

Query Match 6.88; Score 173; DB 2; Length 4000;
Best Local Similarity 51.18; Pred. No. 1.5e-23;
Matches 497; Conservative 0; Mismatches 455; Indels 21; Gaps 3;

```

```

Qy 79 CAGCACTTAGATGAGAGGAGTACTGTGCGGGGCTGTGCTCCCTGGCGAGGCC 138
Db 110 CAGCAGCGCCCAAGAGAGGAGATTTCTTGTGGGGGCTTCTTATTCATTTTGA 169
Qy 139 GAGGAGGCT---GGCTCGCGCAGCGGAGCGCCAGCAGCCCTGTGTGACCGAGTTC 195
Db 170 GTAGAGCTAAAGATCAAGATCTCAATCAAGCGCGAGCTGTGGAATGTATCGAT 229
Qy 196 TCTCAAGGCGCTGCTGTGCGCAGTGTGCGCAATGAATGCGGTGAGAGATCAAC 255
Db 230 AATTTCGCTGGGTTCGCTGTGAGCTATGATATTTGCGATGAGAGATTAACAGC 289
Qy 256 AAGTGGATGTGCTGCGCGGCGCTGAGGCTAGCAGCTTTTGTATGATGCTCGAG 315
Db 290 AGCCAGCGCTTCTTCCCACTTACGCTGGGATACAGATATTTGACACTTGCACAC 349
Qy 316 CCGTGTGCGCATGAACCCAGCCGCTCATGTCTGCGC-----CAAGGACAGGACGCC 369
Db 350 GTTCTAAGCGCTTGAAGCCACCTGAGTTTGTGTCTCAAAATAATGATTTTGTG 409
Qy 370 GACATCGCGCGCTACTGTCACTACAGCAGTACAGCCCGCTGTGCTGTCACTCGG 429
Db 410 AACCTGTAGTGTCTGTCACTGTCAAGCAGCATTTCCCTAGATGTGTGTGTGGA 469
Qy 430 CCCCAGCTGTGAGAGCTGCGCATGTGTCACCGGAGTCTTCTGATGCTGCGC 489
Db 470 GCAACTGAGTAAAGGCTTCCACAGCGCAGTGTGCAATGTGTGTGTGTGTGTG 529
Qy 490 CAGGTACAGTACAGGCTGTAGCATGAGTGTGTGAGCGCGCGGAGACCTTCCCTTC 549
Db 530 CAGGTACATTAATGCTCTCCAGCAGACTCTTCAAGCAAGATCAATTAATGCTTTC 589
Qy 550 TTCCGACACGTCGCGCAGCAGCTGTGTGAGTGTGAGCGCGCGGAGAGCTGTGAGAG 609

```

Db 590 CTCGGAACATCCCAATGATGAGACACAGCCACTGCTCATGCGAGCATCATCATGATAT 649
QY 610 TTGGGTGGAATGAGTGGGCGCCCTGGGCGAGAGACGATGAGCGGCGGAGGCGCTG 669
Db 650 TTCCGCTGGAATGAGTGGGCGCACATTTGACACTATGCGATGAGGCGGCGGATTT 709
QY 670 AGCATCTTCGCGCCCGCGCGGCGAGCGGCGATTCGATGCGGCGAGAGGCGCTG 729
Db 710 GAGAAATTCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 769
QY 730 CCGGTGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 789
Db 770 TCCGAGATCTGATGA-----GGAGAGATCCGATGCTGCTGAGAGGATAT 817
QY 790 AACGAGAGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 849
Db 818 CAAATTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 877
QY 850 TTCACTACAGCATGACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 909
Db 878 ATCAAGAGATGTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 937
QY 910 CTGACCTGACGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCG 969
Db 938 GCGAGCTCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 997
QY 970 TTCTCCGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1029
Db 998 TTGCTCTGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1057
QY 1030 CTGGCGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1042
Db 1058 AGGAGCTCTGTC 1070

RESULT 13

US-08-485-588-2
Sequence 2, Application US/08485588

Patent No. 5688938
GENERAL INFORMATION:

APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: Forrest H. Fuller
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA

ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,588
FILING DATE: 7 June, 1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9

APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994

APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994

APPLICATION NUMBER: U.S. 08/292,827

FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Hebert, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/005
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ. ID NO. 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5006 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
NAME/KEY: CDS
LOCATION: 436..3699
OTHER INFORMATION:
US-08-485-588-2

Query Match 6.88; Score 173; DB 1; Length 5006;

Best Local Similarity 51.18; Pred. No. 1,6e-23; Matches 497; Conservative 0; Mismatches 455; Indels 21; Gaps 3;

QY 79 CAGCACTAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 138
Db 505 CAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 564
QY 139 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 195
Db 565 GTAGCACTAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCA 624
QY 196 TCCCAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 255
Db 625 AATTCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 684
QY 256 AAGTGGATCTGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 315
Db 685 AGCCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 744
QY 316 CCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 369
Db 745 GTTCTAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 804
QY 370 GACATGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 429
Db 805 AACCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 864
QY 430 CCCACAGTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 489
Db 865 GCAAGTGGCTCAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 924
QY 490 CAGGTAGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 549
Db 925 CAGGTAGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 984
QY 550 TTCCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 609
Db 985 CTCGGAACATCCCAATGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 1044

```

OY 610 TTCCGCTGAACTGGGTGGCCGCTGGGACAGCAGAGTACGGCCGAGGGCCCTG 669
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1045 TTCCGCTGAACTGGGTGGCCGCTGGGACAGTGTAGAGTATGGGCGGCGGAGTT 1104
OY 670 AGCATCTTCTGGCCCTGGCCGCGACGCGCATCTGCATCGCCACAGAGGGCTGGTG 729
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1105 GAGAAATTCGAGAGAGAGCTAGAGAAAGGATATCTGCATCGACTTCAGTAACTCATC 1164
OY 730 CCGCTGCCCCGCGGAGTACTCGCGGCTGGGAGAGGTGCAGAGCTCTGCACCAAGGTG 789
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1165 TCCCAGTACTGTAGTA-----GGAAAGATTCAGCATGTGTAGAGTGTAT 1212
OY 790 AACGAGAGAGCGTGCAGGTGTGCTGCTCGCTCCGTCAGCGCCGCGCCAGCCCTC 849
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1213 CAAATTCACAGGCGCAAAATCATCTGTGTTTCTCAGTGGCCGCAATCTTGAGCCCTC 1272
OY 850 TTCAACTACAGCATCAGCAGCAGCTCTCGCCCAAGGTGTGGGTGGCCAGCGAGCTGTG 909
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1273 ATCAAGAGAGATGTCCGGGCAATATCAAGGCAAGATCTGGCTGGCCAGGAGGCTGTG 1332
OY 910 CTGACCTGTGACCTGTGTATGGGGCTGCCCGCATGCGCCAGATGGCAGGCTGTGGC 969
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1333 GCCAGCTCTCCCTGATGCCATGCTCAGTACTTCACAGTGTGGCGGCAACCATTTGGA 1392
OY 970 TTCTCCAGAGGGGTGCCAGCTGCAGAGTTCCCCAGTACGTAGAGAGCGACCTGGCC 1029
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1393 TTGGCTGGAAGGCTGGGGCAGATCCAGGCTTCCGGGAATTCCTGAAGAGGCTCATCCC 1452
OY 1030 CTGGCCACCGACC 1042
    ||| ||| |||
Db 1453 AGGAGTCTGTCC 1465
    ||| ||| |||

```

RESULT 14
US-08-484-565-2
Sequence 2, Application US/08484565
Patent No. 5763569

GENERAL INFORMATION:

APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: James E. Garbett, Jr.
TITLE OF INVENTION: CALCULON RECEPTOR-ACTIVE
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,565
FILING DATE: 7 June, 1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9

APPLICATION NUMBER: 08/353,784

FILING DATE: 9 December, 1994

APPLICATION NUMBER: PCT/US/94/12117

FILING DATE: 21 October, 1994

APPLICATION NUMBER: U.S. 08/292,827

FILING DATE: 23 August, 1994

APPLICATION NUMBER: U.S. 08/141,248

```

FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/006
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5006 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 436..3699
OTHER INFORMATION:
US-08-484-565-2

```

Query Match

6.8%; Score 173; DB 1; Length 5006;

Best local Similarity 51.1%; Pred. No. 1.6e-23;

Matches 497; Conservative 0; Mismatches 455; Indels 21; Gaps 3;

```

OY 79 CAGCACTTAGATGAGGGGAGCTACGTGCTGGGGGGGCTGTTCCTCCCTGGCGAGGCC 138
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 505 CAGCAGGCCCAAGAGAGGGGACATTTATCTGGGGGGCTTCTTATTCATTTTGA 564
OY 139 GAGGAGCT--GGCTCCGACGCGGACGCGCCAGCAGCCCTGTGTACACAGGTTG 195
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 565 GTAGAGAGTAAAGATCAGATCTCAATCAAGCCGAGTCTGTGAATGTATCAGGTAT 624
OY 196 TCCTAAAGGCTCTGTGGGCACTGGCCATGAAAGGCGCGTGAAGATCAACAC 255
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 625 AATTCCTGGTTCGTGCTGTACAGGCTATGATTTGCTCATAGAGAGATTAACAG 684
OY 256 AAGTCGATCTGCTCCGCGGCTGGGCTGAGCAGCTCTTGTATACGTCTGGAG 315
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 685 AGCCAGCCCTTCTCCCACTTGAAGCTGGGATACAGATATTTGACACTTGCAACAC 744
OY 316 CCTGTGTGGCCATGAGCCCAAGCCTCATGTTCTGGC-----CAAGCAGGACGCCG 369
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 745 GTTCTAAGGCTTGAAGCCACCTGAGTTTGTGCTCAAAACAAATATTTCTTTG 804
OY 370 GACATCGCGGCTACTGCAATCAGACGATACGACGACGCGCGCTGTCTGTCTATCGGG 429
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 805 AACCTTGATGATGCTGCAAGCTGCAAGCAATTCCTCTACGATGTGTGTGGGA 864
OY 430 CCCCAGTCAAGAGCTGCCATGATGTCACCGGCAAGTCTTCAAGTCTTCTCTATGCGC 489
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 865 GCAATGCTGTCAGGCTGTCCACGAGCATGGAATTCGTGCGGCTCTTCTACATTTCC 924
OY 925 CAGTCACTTATGCTCTCTCCAGCAGACTCTCACAACAAGAAATCAATCAAGCTTTC 984
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 550 TTCCGACAGCGGCGGAGGACCGTGTGACGTAAGGCGCGCGGAGCACTGTGAGAG 609
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 985 CTCGCAACATCCCAATGATGAGCACCAGGCGCATGCGAGCATCTCAGATAT 1044
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
OY 610 TTCCGCTGAACTGGGTGGCCGCTGGGACAGCAGAGTACGGCCGAGGGGCTGT 669
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 1045 TTCCGCTGGAAGTGGGGGCAATTCACACCTGATGACGATATGGCGCGCGGATTT 1104
QY 670 AGCATCTTCGCGCCCTGCGCGGCAAGGCGCATTCGATCGGCGCAAGAGGCTGGTG 729
Db 1105 GAGAAATTCGAGAGGAGCTGAGAGAAAGGATATCTGATCGACTTGTGAATCATTC 1164
QY 730 CCGCTGCGCGCTGCGGATGACTTCGCGGCGGAGGATGACGATCTCTGACCAAGTG 789
Db 1165 TCCAGTACTGTATGA-----GGAAGAGATCCAGCATGTGTGTAGAGGTGAT 1212
QY 790 AACGAGACGAGCTGACAGGTGTCTGTCTGCTTCGCTTCGCGACGCGCCGACCTC 849
Db 1213 CAAATTCGACGCGCAAGTATGTGTGTTTTCACAGTGGCCCAAGATCTTGACCCCTC 1272
QY 850 TTCAACTACAGCATCAGACGAGGCTCGCGCCCAAGGATGGGGTGGCGAGGCGCTCG 909
Db 1273 ATCAAGAGATATTCGCGCGCAATATCACGCGCAAGATCTGCTGGCGGAGGCTCG 1332
QY 910 CTGACCTCTGACCTGTGATGAGGCTGCGCGCATGGCGGCAAGATGGGACGCTTGGC 969
Db 1333 GCCAGCTCTGCTGATGATGCGCATGCTCAGTACTTCCACGATGGTGGCGGACCATTTGA 1392
QY 970 TTCTCCAGAGGGGTGCCCACTGACAGAGTTCCTCCAGTACGTGAAGACCACTGGCC 1029
Db 1393 TTGCTCTGAGAGGCTGGGCGAGATCCAGGCTTCGCGGAATTCCTGAAGAGGCTCATCC 1452
QY 1030 CTGGCCACCGAC 1042
Db 1453 AGGAAGTCTGTCC 1465

```

RESULT 15
US-08-480-751-2
Sequence 2, Application US/08480751
Patent No. 585684

GENERAL INFORMATION:
APPLICANT: Edward F. Nemeeth
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: Forrest H. Fuller
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,751
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/35,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248

```

FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Hebert, Sheldon O.
REGISTRATION NUMBER: 38,179
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5006 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
FEATURE:
NAME/KEY: CDS
LOCATION: 436..3699
OTHER INFORMATION:
US-08-480-751-2

```

Query Match 6.8%; Score 173; DB 2; Length 5006;
Best Local Similarity 51.1%; Pred. No. 1.6e-23;
Matches 497; Conservative 0; Mismatches 455; Indels 21; Gaps 3;

```

QY 79 CAGCACTTAGATGAGAGGGGACTACGTGCTGGGGGCTGTTCCTCCGCGGAGGCC 138
Db 505 CAGCGAGCCCAAGAGAGGAGGACATATCTTGGGGGCTTCTTCTATTCATTTTGA 564
QY 139 GAGGAGCT---GGCCCTCCGAGCGGAGGAGCGGCGGAGGAGGAGGAGGAGGAGG 195
Db 565 GTACAGCTAAGATCAAGATCTAAATCAAGCGGAGGAGGAGGAGGAGGAGGAGG 624
QY 196 TCCTCAAGCGGCTGCTGGGCACTGCGCATGAAATGCGCGGAGGAGGAGGAGGAGG 255
Db 625 AATTCGCTGGGCTTCTGCTGTTACAGGCTATGATTTTCCAGAGGAGGAGGAGGAG 684
QY 256 AAGTGGATGCTGCGCGGCGGCTGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGG 315
Db 685 AGCCAGCGCTTCTTCCCACTTACGCTGGGATACAGATATTTGACACTTGCACACC 744
QY 316 CCGTGGTGGCGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 369
Db 745 GTTCTAAGGCTTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 804
QY 370 GAGATCGCGGCTTCACTCACTACAGCAGTACAGGCGCGGCTGCTGCTGATCGGG 429
Db 805 AACCTTATGATGCTCTCACTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 864
QY 430 CCCCAGCTGACAGCTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 489
Db 865 GCAACTGGCTCAGGCGCTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 924
QY 490 CAGGTGAGTACGCTGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 549
Db 925 CAGGTGAGTACGCTGCTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 984
QY 550 TTGCGACGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 609
Db 985 CTGCGAGCAGTCCCAATGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAT 1044
QY 610 TTGCGCTGGAAGTGGTGGCGCGGCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 669

```

```
Db 1045 TTCCGCTGGAACTGGGTGGGCGCAATTGCAGCTGATGAGACTATGCGCGCGCGGGGAAFT 1104
QY 670 AGCATCTTCTGCGCCCTGGCCGCGGACGCGGCAATGCAATCGCGGACAGAGGCTGCTG 729
Db 1105 GAGAAATTCGAGAGAGAGAGAGAGAGAGATATCTGATCGACTTCAGTGAATCATC 1164
QY 730 CCGCTGCCCGCGGATGACTCGCGGCTGGGGAGAGTGCAGAGCTCTGACACAGTG 789
Db 1165 TCCGATCTGTGATGA-----GGAGAGATCCAGCATGTGTAGAGGTGATT 1212
QY 790 AACGAGAGAGCGGTGACAGGTGCTGCTGCTTCCGCTCCGTGACAGCGCGCCACGCGCTC 849
Db 1213 CAAAATTCAGGCGCAAGATCATGCTGTTTCTCCAGTGGCCGACATCTTGAAGCCCTC 1272
QY 850 TTCAACTACAGCATGACGACAGGCTCTGCCCCAAGGTGTGGGTGGCCAGCGAGGCTTG 909
Db 1273 ATCAAGGAGATGTCTCGGCGCAATATCACGGGCAAGATCTGGCTGGCCAGGAGGCTTG 1332
QY 910 CTGACCTCTGACCTGTGATGAGGCTGCGCGGATGCGGCAATGAGGACAGGTGCTTGGC 969
Db 1333 GCCAGCTCTCTCCTGATGCGCATGCGCTCAGTACTTCCAGGTGTTGGCGCACCATTTGA 1392
QY 970 TTCTTCAGAGGAGTGCAGCGAGTTCCTCCAGTACGTGAAGAGCGACCTTGCGC 1029
Db 1393 TTCGCTTGAAGGCTGGGCGAGATCCAGGCTTCGGGGAATTCCTGAAGAGGTCCATCCC 1452
QY 1030 CTGGCCACCGACC 1042
Db 1453 AGGAAGTCTGTCC 1465
```

Search completed: May 23, 2003, 18:24:46
Job time : 153.053 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4_P5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 23, 2003, 12:46:24 ; Search time 315.884 Seconds
(Without alignments)
10697.149 Million cell updates/sec

Title: US-09-927-315-14

Perfect score: 2559

Sequence: 1 atgtctggccctgtctgtctc.....atcagggaacatgagtga 2559

Scoring table:

Gapop 10.0 , Gapext 1.0

Searched: 828747 seqs, 660231138 residues

Total number of hits satisfying chosen parameters: 1657494

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published_Applications_NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PC1_NEM_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEM_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEM_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEM_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09_NEM_PUB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US10_NEM_PUB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US60_NEM_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2559	100.0	2559	US-09-927-315-14	Sequence 14, Appl
2	2559	100.0	2559	US-10-035-045-3	Sequence 3, Appl
3	2557.4	99.9	2559	US-09-897-427A-5	Sequence 5, Appl
4	2557.4	99.9	2559	US-09-897-427A-7	Sequence 7, Appl
5	2557.4	99.9	3458	US-10-261-482-1	Sequence 1, Appl
6	2555.8	99.9	2559	US-10-282-837-15	Sequence 15, Appl
7	2555.8	99.9	3489	US-09-796-338A-13	Sequence 13, Appl
8	2555.8	99.9	3489	US-09-796-338A-13	Sequence 13, Appl
9	2553	99.1	2553	US-09-799-629-3	Sequence 3, Appl
10	2537	99.1	3200	US-09-927-315-13	Sequence 13, Appl
11	1947.4	76.1	8001	US-10-261-482-3	Sequence 3, Appl
12	1897	74.1	3563	US-09-799-629-20	Sequence 20, Appl
13	1897	74.1	3563	US-10-035-045-20	Sequence 2, Appl
14	1666	65.1	2687	US-09-799-629-2	Sequence 2, Appl
15	1666	65.1	2687	US-10-035-045-2	Sequence 2, Appl
16	1470.8	57.5	2577	US-09-799-629-13	Sequence 13, Appl
17	1470.8	57.5	2577	US-09-927-315-24	Sequence 24, Appl
18	1470.8	57.5	2577	US-10-035-045-13	Sequence 13, Appl
19	1470.8	57.5	2577	US-10-035-045-13	Sequence 13, Appl

20	1456.4	56.9	2577	US-09-927-315-17	Sequence 17, Appl
21	1454.8	56.9	2577	US-09-927-315-22	Sequence 22, Appl
22	1450	56.7	2577	US-09-927-315-19	Sequence 19, Appl
23	873.2	34.1	3240	US-09-927-315-16	Sequence 16, Appl
24	871.6	34.1	3200	US-09-927-315-21	Sequence 21, Appl
25	381	14.9	876	US-09-799-629-1	Sequence 1, Appl
26	381	14.9	876	US-10-035-045-1	Sequence 1, Appl
27	378.8	14.8	558	US-09-799-629-11	Sequence 11, Appl
28	378.8	14.8	558	US-10-035-045-11	Sequence 11, Appl
29	366	14.3	2526	US-09-819-946-1	Sequence 1, Appl
30	366	14.3	2951	US-09-819-946-9	Sequence 9, Appl
31	365.6	14.3	2526	US-09-897-427A-9	Sequence 1, Appl
32	365.6	14.3	2526	US-09-799-629-16	Sequence 16, Appl
33	365.6	14.3	2526	US-10-035-045-16	Sequence 16, Appl
34	357	14.0	2292	US-09-819-946-3	Sequence 3, Appl
35	350.8	13.7	2771	US-09-361-652-4	Sequence 4, Appl
36	350.8	13.7	2771	US-09-927-315-4	Sequence 4, Appl
37	346.6	13.5	2333	US-09-361-652-6	Sequence 6, Appl
38	346.6	13.5	2333	US-09-927-315-6	Sequence 6, Appl
39	344.4	13.5	2579	US-09-361-652-5	Sequence 5, Appl
40	344.4	13.5	2579	US-09-927-315-5	Sequence 5, Appl
41	319.6	12.5	552	US-09-799-629-9	Sequence 9, Appl
42	319.6	12.5	552	US-10-035-045-9	Sequence 9, Appl
43	296.8	11.6	2520	US-09-897-427A-3	Sequence 3, Appl
44	296.8	11.6	2520	US-10-035-045-23	Sequence 23, Appl
45	284	11.1	2532	US-10-096-144-4	Sequence 4, Appl

ALIGNMENTS

RESULT 1

US-09-927-315-14
Sequence 14, Application US/09927315
Publication No. US20030040045A1
GENERAL INFORMATION:
APPLICANT: Zuker, Charles S.
APPLICANT: Ryba, Nicholas J.P.
APPLICANT: Nelson, Greg
APPLICANT: Hoon, Mark A.
APPLICANT: Chandrasekar, Jayaram
APPLICANT: Zhang, Yifeng
APPLICANT: The Regents of the University of California
APPLICANT: The Government of the United States of America
APPLICANT: as represented by the Secretary of the
Department of Health and Human Services
TITLE OF INVENTION: Mammalian Sweet Taste Receptors
FILE REFERENCE: 02307E-120110US
CURRENT APPLICATION NUMBER: US/09/927,315
CURRENT FILING DATE: 2001-08-10
PRIOR APPLICATION NUMBER: US 60/302,898
PRIOR FILING DATE: 2001-07-03
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 14
LENGTH: 2559
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: human T1R3 sweet taste receptor CDS
US-09-927-315-14

Query Match 100.0% Score 2559; DB 9; Length 2559;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGCTGGCCCTGCTGCTGGCCCTGACCCCTGCGACCCCTGGAGCGGG 60
DB 1 ATGCTGGCCCTGCTGCTGGCCCTGACCCCTGCGACCCCTGGAGCGGG 60
QY 61 GCCCATTTGCTGCTGCTGACAGCACTTAGGATGAGGAGGACTGCTGGGGGGCTG 120
DB 61 GCCCATTTGCTGCTGCTGACAGCACTTAGGATGAGGAGGAGGAGGAGGCTG 120

QY 121 TTCCCTGGGCGAGGCGGAGGAGGCTGCTCCGACGCGGACAGGCGCCAGGACCTT 180
Db 121 TTCCCTGGGCGAGGCGGAGGAGGCTGCTCCGACGCGGACAGGCGCCAGGACCTT 180
QY 181 GTGTGACACAGGTTCTCTCAACGCGCTGCTGTGGGCACTGAGCATGAAATGCGCTG 240
Db 181 GTGTGACACAGGTTCTCTCAACGCGCTGCTGTGGGCACTGAGCATGAAATGCGCTG 240
QY 241 GAGGAGATCAACAACTGAGATGCTGCGCGGCGGCTGCGGCTGAGCACTGCTT 300
Db 241 GAGGAGATCAACAACTGAGATGCTGCGCGGCGGCTGCGGCTGAGCACTGCTT 300
QY 301 GATAGTGTCTGAGGAGCTGTGTGGCCATGAAGCCGACCTCATGTTCTCTGCGCAAGCA 360
Db 301 GATAGTGTCTGAGGAGCTGTGTGGCCATGAAGCCGACCTCATGTTCTCTGCGCAAGCA 360
QY 361 GGCAGCGCGGACATGCGCGGCTACTGCACTACAGCAGTACAGAGCCCGGTGTGCTGCT 420
Db 361 GGCAGCGCGGACATGCGCGGCTACTGCACTACAGCAGTACAGAGCCCGGTGTGCTGCT 420
QY 421 GTCATCGGCGGCGGCTGCTGAGAGCTGCGCATGAGTACCGGCAAGTCTTCACTGCTTTC 480
Db 421 GTCATCGGCGGCGGCTGCTGAGAGCTGCGCATGAGTACCGGCAAGTCTTCACTGCTTTC 480
QY 481 CTGATGCGGCGGAGTACGCTGAGTACGATGAGTGTGAGCGCGCGGAGACCTTC 540
Db 481 CTGATGCGGCGGAGTACGCTGAGTACGATGAGTGTGAGCGCGCGGAGACCTTC 540
QY 541 CCCGCTCTCTCCGACCGGCTGCGGACCGGCTGCGGAGCTGAGCGAGCGCGCGGAGCTG 600
Db 541 CCCGCTCTCTCCGACCGGCTGCGGACCGGCTGCGGAGCTGAGCGAGCGCGCGGAGCTG 600
QY 601 CTGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 660
Db 601 CTGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 660
QY 661 CAGGAGCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 720
Db 661 CAGGAGCTGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 720
QY 721 GGCCTGGGCGGCTGCGGCGGCTGCGGAGCTGCGGCGGCTGCGGAGCTGCTG 780
Db 721 GGCCTGGGCGGCTGCGGCGGCTGCGGAGCTGCGGCGGCTGCGGAGCTGCTG 780
QY 781 CACCAAGTGAACAGAGAGGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 840
Db 781 CACCAAGTGAACAGAGAGGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 840
QY 841 CAGGCGCTCTCAACATACAGATACAGAGAGGCTGCGGCGGCTGCGGAGCTGCTG 900
Db 841 CAGGCGCTCTCAACATACAGATACAGAGAGGCTGCGGCGGCTGCGGAGCTGCTG 900
QY 901 GAGGCGTGGGCTGAGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 960
Db 901 GAGGCGTGGGCTGAGCTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 960
QY 961 GTGCTGGCTGTCTCAAGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1020
Db 961 GTGCTGGCTGTCTCAAGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1020
QY 1021 CACCTGGCGCTGCGGAGCGGCGCTTCTGCTGCTGCGGCTGCGGAGAGGAGGAGGCT 1080
Db 1021 CACCTGGCGCTGCGGAGCGGCGCTTCTGCTGCTGCGGCTGCGGAGAGGAGGAGGCT 1080
QY 1081 CTGAGAGGAGAGCTGAGGAGGAGGAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAG 1140
Db 1081 CTGAGAGGAGAGCTGAGGAGGAGGAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAG 1140
QY 1141 GTGAGCGCAGGAGCTAAATCAACAGAGCTTCTGCTGAGAGCTGTATAGGCTG 1200
Db 1141 GTGAGCGCAGGAGCTAAATCAACAGAGCTTCTGCTGAGAGCTGTATAGGCTG 1200

QY 1201 GCCAGGCGCTGACACACACTTCTAGTGAACGCGCTGAGGCTGCGCGGAGGACCC 1260
Db 1201 GCCAGGCGCTGACACACACTTCTAGTGAACGCGCTGAGGCTGCGCGGAGGACCC 1260
QY 1261 GTGAAGCGCTGAGAGCTCTGAGAGACATGTACACCTGACCTTCCAGCTGAGGCTG 1320
Db 1261 GTGAAGCGCTGAGAGCTCTGAGAGACATGTACACCTGACCTTCCAGCTGAGGCTG 1320
QY 1321 CCGGTGCGGCTGAGAGAGGAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1380
Db 1321 CCGGTGCGGCTGAGAGAGGAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 1380
QY 1381 TGCGAGGAGCTGAGTGTGAGGAGGAGGAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAG 1440
Db 1381 TGCGAGGAGCTGAGTGTGAGGAGGAGGAGGAGGAGTGTGAGTGTGAGTGTGAGTGTGAG 1440
QY 1441 GAGGCGCTGAAAGATCCGCTGAGCAGTGTGAGCAGTGTGAGCAGTGTGAGCAGTGTGAG 1500
Db 1441 GAGGCGCTGAAAGATCCGCTGAGCAGTGTGAGCAGTGTGAGCAGTGTGAGCAGTGTGAG 1500
QY 1501 CGGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
Db 1501 CGGAGTGTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560
QY 1561 TGTGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
Db 1561 TGTGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620
QY 1621 TGTGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
Db 1621 TGTGTGAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1680
QY 1681 TTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
Db 1681 TTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1740
QY 1741 GGCCTGTGTGCTGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
Db 1741 GGCCTGTGTGCTGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800
QY 1801 GGCCTGTGTGCTGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1860
Db 1801 GGCCTGTGTGCTGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1860
QY 1861 GTCTGTCTGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1920
Db 1861 GTCTGTCTGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1920
QY 1921 CACCTCCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1980
Db 1921 CACCTCCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1980
QY 1981 GAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2040
Db 1981 GAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2040
QY 2041 GGCCTGTGTGCTGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2100
Db 2041 GGCCTGTGTGCTGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2100
QY 2101 GTGAGCTTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2160
Db 2101 GTGAGCTTCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2160
QY 2161 CACTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2220
Db 2161 CACTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2220
QY 2221 GGCCTTGTGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2280
Db 2221 GGCCTTGTGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2280
QY 2281 GCCCGGCGCTGACCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2340

Db 2281 GCGCGGGGCTACCTTGGCCATGCTGGGCTACTTCACTACCTAGCTGGGCTCTTCTTGTGGCC 2340
QY 2341 CTCCTGGCCAAATGTGAGGTGTCTCTCAGGCCCCGCGCTGACATGGGCGCCCTCTGCTC 2400
Db 2341 CTCCTGGCCAAATGTGAGGTGTCTCTCAGGCCCCGCGCTGACATGGGCGCCCTCTGCTC 2400
QY 2401 TGTGTCTGGGATGCTGGCTGCTCTCCACCTGCGGAGGTGTACTGCTCATGGGGCAG 2460
Db 2401 TGTGTCTGGGATGCTGGCTGCTCTCCACCTGCGGAGGTGTACTGCTCATGGGGCAG 2460
QY 2461 CCAGGGCTCAACACCCCGAGTCTCTCAGGAGGGGCGCTGGGAGATGCCCAAGGCCAG 2520
Db 2461 CCAGGGCTCAACACCCCGAGTCTCTCAGGAGGGGCGCTGGGAGATGCCCAAGGCCAG 2520
QY 2521 AATGACGGGAACACAGGAATCAGGGGGAACATGAGTGA 2559
Db 2521 AATGACGGGAACACAGGAATCAGGGGGAACATGAGTGA 2559

RESULT 2

US-10-035-045-3
Sequence 3, Application US/10035045
Publication No. US2003005448A1
GENERAL INFORMATION:
APPLICANT: ADLER, JON ELLIOT
APPLICANT: LI, XIADONG
APPLICANT: STASZEWSKI, LENA
APPLICANT: O'CONNELL, SHAWN
APPLICANT: ZOZULYA, SERGEY
TITLE OF INVENTION: TLR TASTE RECEPTORS AND GENES ENCODING SAME
FILE REFERENCE: 078003-0280681
CURRENT APPLICATION NUMBER: US/10/035,045
CURRENT FILING DATE: 2002-01-03
PRIOR APPLICATION NUMBER: 60/259,227
PRIOR FILING DATE: 2001-01-03
PRIOR APPLICATION NUMBER: 60/284,547
PRIOR FILING DATE: 2001-04-19
NUMBER OF SEQ ID NOS: 24
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3
LENGTH: 2559
TYPE: DNA
ORGANISM: Homo sapiens
US-10-035-045-3

Query Match 100.0%; Score 2559; DB 9; Length 2559;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2559; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 AAGCTGGGCGGCTGCTCTGAGGCGCTCAGGCGCTGCTGCTGACCTGGAGCGGG 60
Db 1 AAGCTGGGCGGCTGCTCTGAGGCGCTCAGGCGCTGCTGCTGACCTGGAGCGGG 60
QY 61 GCCCATTTGCTGCTGACAGCACTTAGATGAAGGGAGTACGTCTGGGGGGCTG 120
Db 61 GCCCATTTGCTGCTGACAGCACTTAGATGAAGGGAGTACGTCTGGGGGGGGCTG 120
QY 121 TTCCCTGAGGAGGCGGAGGAGGCTGCTCCGACGCGGACAGCGGCCAGACCCCT 180
Db 121 TTCCCTGAGGAGGCGGAGGAGGCTGCTCCGACGCGGACAGCGGCCAGACCCCT 180
QY 181 GTGTGACACAGGTCTCTCCCAACGGCCCTGCTGGGCACTGGCCATGAAATGGCCGTG 240
Db 181 GTGTGACACAGGTCTCTCCCAACGGCCCTGCTGGGCACTGGCCATGAAATGGCCGTG 240
QY 241 GAGAGATCAACAACAGTGTGATGCTGCTGCGGGGCTGCGCTGGGCTACGACTTT 300
Db 241 GAGAGATCAACAACAGTGTGATGCTGCTGCGGGGCTGCGCTGGGCTACGACTTT 300
QY 301 GATACGTCTCGAGGCTGTGTGGCTGAAGAGCCAGGCTCATGTTCTTGCCCAAGGCA 360
Db 301 GATACGTCTCGAGGCTGTGTGGCTGAAGAGCCAGGCTCATGTTCTTGCCCAAGGCA 360

QY 361 GGCAGCGCGAGATAGCGCGCTTACAGCACTACAGCACTACAGCGCGCGCTGTGCT 420
Db 361 GGCAGCGCGAGATAGCGCGCTTACAGCACTACAGCACTACAGCGCGCGCTGTGCT 420
QY 421 GTCATCGGGCCCACTGTCAGAGTTCGCAAGTGCATGAGTGCAGGAAATTTCTGACTTTC 480
Db 421 GTCATCGGGCCCACTGTCAGAGTTCGCAAGTGCATGAGTGCAGGAAATTTCTGACTTTC 480
QY 481 CTCATGCCCCAGGTACGTACGCTGTAGCATGAGTGTCTGAGCGCCCGGAGACCTTC 540
Db 481 CTCATGCCCCAGGTACGTACGCTGTAGCATGAGTGTCTGAGCGCCCGGAGACCTTC 540
QY 541 CCTCTCTTCTTCCGACCGCTGCGCCAGCGACGCTGTGAGCTGACGCGCGCGGAGCTG 600
Db 541 CCTCTCTTCTTCCGACCGCTGCGCCAGCGACGCTGTGAGCTGACGCGCGCGGAGCTG 600
QY 601 CTGCAAGAGTTGCGCTGGAATGAGGTGGCGCCCTTGGCGAGCAGCAGTACGGCCGG 660
Db 601 CTGCAAGAGTTGCGCTGGAATGAGGTGGCGCCCTTGGCGAGCAGCAGTACGGCCGG 660
QY 661 CAGGGCTGAGCATCTTCTGCGCCCTGCGCGGCAAGCGGCACTGCAATCGCCACAG 720
Db 661 CAGGGCTGAGCATCTTCTGCGCCCTGCGCGGCAAGCGGCACTGCAATCGCCACAG 720
QY 721 GGGCTGTGGCGGCTGCGCCCGGCTGCGAGTGTGCGGCTGGGGAAGGTGAGGAGCTTC 780
Db 721 GGGCTGTGGCGGCTGCGCCCGGCTGCGAGTGTGCGGCTGGGGAAGGTGAGGAGCTTC 780
QY 781 CACAGGTGAACACAGAGCAGTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
Db 781 CACAGGTGAACACAGAGCAGTGTGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 840
QY 841 CAGCGCTTCTTCACTACAGCATCAGCAGCAGAGGCTTCTGCGCCCAAGGTGTGGTGGC 900
Db 841 CAGCGCTTCTTCACTACAGCATCAGCAGCAGAGGCTTCTGCGCCCAAGGTGTGGTGGC 900
QY 901 GAGGCGTGGCAGCTGCTGACCTGCTGATGAGTGGGCGCGGCGGATGGCCAGATGGGCG 960
Db 901 GAGGCGTGGCAGCTGCTGACCTGCTGATGAGTGGGCGCGGCGGATGGCCAGATGGGCG 960
QY 961 GTGCTTGGCTTCTCAGAGAGGGGTGCCAGCTGCAAGATTTCCCAAGTACGTGAAGAG 1020
Db 961 GTGCTTGGCTTCTCAGAGAGGGGTGCCAGCTGCAAGATTTCCCAAGTACGTGAAGAG 1020
QY 1021 CACCTGGCCCTGCGCAGCAGCAGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
Db 1021 CACCTGGCCCTGCGCAGCAGCAGCGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
QY 1081 CTGAGAGAGAGGTGTGGGCGAGGCTGCGGCAAGTGTGATGATGATGATGATGATG 1140
Db 1081 CTGAGAGAGAGGTGTGGGCGAGGCTGCGGCAAGTGTGATGATGATGATGATGATG 1140
QY 1141 GTGAGGCGAGGGCTTAATACCAACAGAGTCTGCTGTATGAGCAGCTGTATAGCGTG 1200
Db 1141 GTGAGGCGAGGGCTTAATACCAACAGAGTCTGCTGTATGAGCAGCTGTATAGCGTG 1200
QY 1201 GCCCAGGCGCTGACAACTTTCAAGTCAAGCGCTGAGGCTGCGCGGCGGAGACCC 1260
Db 1201 GCCCAGGCGCTGACAACTTTCAAGTCAAGCGCTGAGGCTGCGCGGCGGAGACCC 1260
QY 1261 GTGAGGCGCTGAGCTGCTGAGGAAACATGATCAACCTGATCAAGTGGGCGGCTG 1320
Db 1261 GTGAGGCGCTGAGCTGCTGAGGAAACATGATCAACCTGATCAAGTGGGCGGCTG 1320
QY 1321 CGGCTGCGGTGAGAGCAGCGGAAACGTGAGATGATGATGATGATGATGATGATGATG 1380
Db 1321 CGGCTGCGGTGAGAGCAGCGGAAACGTGAGATGATGATGATGATGATGATGATGATG 1380
QY 1381 TGGCAGGGCTCAGTGGCCAGGCTTCAAGCAGTGTGGGAGGTTCAAGGCGCTCAGGACA 1440
Db 1381 TGGCAGGGCTCAGTGGCCAGGCTTCAAGCAGTGTGGGAGGTTCAAGGCGCTCAGGACA 1440

OY	1441	GAGGCGCTGAAGAAATCCGCTGCGACACAGTCTGACAAACACAAGCCGCTGTCOCGGGCTCG	1500
Db	1441	GAGCGCCTGAAGATCCGCTGCGACACAGTCTGACAAACACAAGCCGCTGTCOCGGGCTCG	1500
OY	1501	CGGACAGTGCAGAGGAGGCGCCAGGTGCGCGCGGATTCAAAGGGTTCTCACTCCGTGCTACAC	1560
Db	1501	CGGACAGTGCAGAGGAGGCGCCAGGTGCGCGCGGATTCAAAGGGTTCTCACTCCGTGCTACAC	1560
OY	1561	TGTGTGACACTGCGAGGCGGGCAGCTACCGGCAAAACCCAGACAGACATGAGCTGCACCTTT	1620
Db	1561	TGTGTGACACTGCGAGGCGGGCAGCTACCGGCAAAACCCAGACAGACATGAGCTGCACCTTT	1620
OY	1621	TGTGGCCAGGATGATGATGTCCCGCGAGCGAAGACACAGCTGTTCCGCGCAGAGTCTGG	1680
Db	1621	TGTGGCCAGGATGATGATGTCCCGCGAGCGAAGACACAGCTGTTCCGCGCAGAGTCTGG	1680
OY	1681	TTTCCGTGAGATGGGCGAGAGCGGCTGATGCTCTCTCTCTGCTGCTGAGCCCTGGGCGTG	1740
Db	1681	TTTCCGTGAGATGGGCGAGAGCGGCTGATGCTCTCTCTCTGCTGCTGAGCCCTGGGCGTG	1740
OY	1741	GGCCTTGTGCTGCTGCTTTTGGGAGTGTTCATTACCATGAGAGACCCACTGGTTCAG	1800
Db	1741	GGCCTTGTGCTGCTGCTTTTGGGAGTGTTCATTACCATGAGAGACCCACTGGTTCAG	1800
OY	1801	GGCTCGGAGGAGGAGCCCGAGGCGCTGGTTTGGGCTGTGCTGCTGGGCGTGGTCTGCTCAGC	1860
Db	1801	GGCTCGGAGGAGGAGCCCGAGGCGCTGGTTTGGGCTGTGCTGCTGGGCGTGGTCTGCTCAGC	1860
OY	1861	GTTCCTCTGTTCCTCTGGGACAGCCAGCCCTCCGATGCTGAGCCAGAGCCCTTTGTC	1920
Db	1861	GTTCCTCTGTTCCTCTGGGACAGCCAGCCCTCCGATGCTGAGCCAGAGCCCTTTGTC	1920
OY	1921	CACCTCCCGGCTCACAGGGCTGCTGAGCACACTTTCCTGCAAGCGCGCAGATCTTGTG	1980
Db	1921	CACCTCCCGGCTCACAGGGCTGCTGAGCACACTTTCCTGCAAGCGCGCAGATCTTGTG	1980
OY	1981	GAGTCAGAACTGACCTCTGAGCTGGGAGAGACCGGCTGAGTGGGCGTGGGGGGGCGCTGG	2040
Db	1981	GAGTCAGAACTGACCTCTGAGCTGGGAGAGACCGGCTGAGTGGGCGTGGGGGGGCGCTGG	2040
OY	2041	GGCTGAGGATGATGATGCTGAGGACATGCTGTGAGAGTGGCACTGAGCACTGGATACCTG	2100
Db	2041	GGCTGAGGATGATGATGCTGAGGACATGCTGTGAGAGTGGCACTGAGCACTGGATACCTG	2100
OY	2101	GTCGCTTCCCGCGGAGGTGTGACGAGACTGGCACATGCTGCCACAGAGGCGCTGGTG	2160
Db	2101	GTCGCTTCCCGCGGAGGTGTGACGAGACTGGCACATGCTGCCACAGAGGCGCTGGTG	2160
OY	2161	CACGCGCGACAGGCTCTGTAGGTCACGTTGGGCGTGTGGGCGACGCGCAAGTGTCCAGCGTG	2220
Db	2161	CACGCGCGACAGGCTCTGTAGGTCACGTTGGGCGTGTGGGCGACGCGCAAGTGTCCAGCGTG	2220
OY	2221	GGCTTTTCTGTGCTCTGAGGACACTTTTCTGTGTGCGAGGACAGCCGGGTGTACAAACGT	2280
Db	2221	GGCTTTTCTGTGCTCTGAGGACACTTTTCTGTGTGCGAGGACAGCCGGGTGTACAAACGT	2280
OY	2281	GGCCGTGAGCCTCACCTTGTGCATGCTGAGGCTACTTCATCACTGGGTGTCTCTTTGTGCC	2340
Db	2281	GGCCGTGAGCCTCACCTTGTGCATGCTGAGGCTACTTCATCACTGGGTGTCTCTTTGTGCC	2340
OY	2341	CTTCCGCGCAATGTGAGGAGGTGTGTCACAGGCGCGGCTGTCAATATGGGGGCGCTCTGTGTC	2400
Db	2341	CTTCCGCGCAATGTGAGGAGGTGTGTCACAGGCGCGGCTGTCAATATGGGGGCGCTCTGTGTC	2400
OY	2401	TGTGTCTGGGACATCTGAGTGCCTTTCACATGCGCAGGATGTTACTGTGATATGGGAG	2460
Db	2401	TGTGTCTGGGACATCTGAGTGCCTTTCACATGCGCAGGATGTTACTGTGATATGGGAG	2460
OY	2461	CCAGAGGCTCAACACCCCGATGTTCTTCTGTGGAGGGGCGCTGGGAGATGCCCAAAGCCAG	2520
Db	2461	CCAGAGGCTCAACACCCCGATGTTCTTCTGTGGAGGGGCGCTGGGAGATGCCCAAAGCCAG	2520
OY	2521	AATGACGGGAAACAGGAATTCAGGGGAAAAATGACTGA	2559

[illegible]

QY	661	CAGGCGCTGAGACAACTTCTTCTGGCCCTTGGCCCGGCGGACAGGGGCGATCTGCATCGCGACAGAG	720
Db	661	CAGGGCTCGAGCACTTCTTCTGGCCCTTGGCCCGGCGGACAGGGGCGATCTGCATCGCGACAGAG	720
QY	721	GGCGTGGTGGCGCTGGCCCGCTGGCGATGATCTCCGCGCTGGGGAAGGTGCAGAGACTCTCTG	780
Db	721	GGCCGTGGTGGCGCGTGGCCCGCTGGCGATGATCTGGCGAGTGGGGGAAGGTGCAGAGACTCTCTG	780
QY	781	CACCAAGTGAACCAAGACAGCGTGCAGAGTGTGTCTGTCTTGTGCTCCGTGCAGCGCGCC	840
Db	781	CACCAAGGTGAACCAAGACAGCGTGCAGAGTGTGTCTGTCTTGTGCTCCGTGCAGCGCGCC	840
QY	841	CACGCGCCCTTCAACTACACATCAGACAGAGAGGCTGTGCCCCAAGTGTGGGTGGCCAGC	900
Db	841	CACGCGCCCTTCAACTACACATCAGACAGAGAGGCTGTGCCCCAAGTGTGGGTGGCCAGC	900
QY	901	GAGGCGCTGGCTGACCTTCTGACCTGTTCATGTGGGCTCTCCCGCATGGCCAGATGGGCGACG	960
Db	901	GAGGCGCTGGCTGACCTTCTGACCTGTTCATGTGGGCTCTCCCGCATGGCCAGATGGGCGACG	960
QY	961	GTGCTTGGCTTCTCTCCAGAGGGGTGCCAGAGTGCACAGATTTCCCCAGTACGTGAAGACG	1020
Db	961	GTGCTTGGCTTCTCTCCAGAGGGGTGCCAGAGTGCACAGATTTCCCCAGTACGTGAAGACG	1020
QY	1021	CACCTGGCGCTTGGGCAACGAGCCCGGCTTCTGTGCTCCCTGGGGGAGAGGGAGCAGAGGT	1080
Db	1021	CACCTGGCGCTTGGGCAACGAGCCCGGCTTCTGTGCTCCCTGGGGGAGAGGGAGCAGAGGT	1080
QY	1081	CTGAGAGAGAGACGTGTGGGGCCAGCGCTGCCCGCAGTGTGATGCATACAGCTGCAGAAC	1140
Db	1081	CTGAGAGAGAGACGTGTGGGGCCAGCGCTGCCCGCAGTGTGATGCATACAGCTGCAGAAC	1140
QY	1141	GTGAGCCGAGGGCTAAATTCACACACAGAGGTTCTGTGTACGACAGCTGTATATGCGTG	1200
Db	1141	GTGAGCCGAGGGCTAAATTCACACACAGAGGTTCTGTGTACGACAGCTGTGTATATGCGTG	1200
QY	1201	GCCAGGCGCTTGCAACAACACTTCTTCAGTGCACAGCCTCAGGCTGCGCCCGCGAGGACCCC	1260
Db	1201	GCCAGGCGCTTGCAACAACACTTCTTCAGTGCACAGCCTCAGGCTGCGCCCGCGAGGACCCC	1260
QY	1261	GTGAAAGCCCTGGGCGCTCCGAGGAACATGTACAACCTGACCTTCCAGCTGGGGGCGCTG	1320
Db	1261	GTGAAAGCCCTTGCGAGCTCTCGAGGAACATGTACAACCTGACCTTCCAGCTGGGGGCGCTG	1320
QY	1321	CCGCTGCGGTTTGCACACAGCAGCGGAAACGTGTGACATGAGATAGACCTGAAGCTGTGGGTG	1380
Db	1321	CCGCTGCGGTTTGCACACAGCAGCGGAAACGTGTGACATGAGATAGACCTGAAGCTGTGGGTG	1380
QY	1381	TGGCAGAGGCTCATGTGCCCAAGGCTTCACAGAGTGTGGAGTGTTCACAGGCAAGCCTCAGAGCA	1440
Db	1381	TGGCAGAGGCTCATGTGCCCAAGGCTTCACAGAGTGTGGAGTGTTCACAGGCAAGCCTCAGAGCA	1440
QY	1441	GAGCGCCGTAAGATCCGCTGGCAACAGTCTGGAACAACAACAACCCCGTGTGCCGGGTGCG	1500
Db	1441	GAGCGCCGTAAGATCCGCTGGCAACAGTCTGGAACAACAACAACCCCGTGTGCCGGGTGCG	1500
QY	1501	CGGCAAGTGCAGAGAGGGCCAGAGTGTGCCCGGGGTTCACATCCGTGCTACAGAC	1560
Db	1501	CGGCAAGTGCAGAGAGGGCCAGAGTGTGCCCGGGGTTCACATCCGTGCTACAGAC	1560
QY	1561	TGTGTGGATCTCGAGGGGGGCGGCGACCTTCACGGCAAAACCCAGAGCATGCGCTGCACCTTT	1620
Db	1561	TGTGTGGATCTCGAGGGGGGCGGCGACCTTCACGGCAAAACCCAGAGCATGCGCTGCACCTTT	1620
QY	1621	TGTGTGGATGATGATGTGTCCCGGAGCGAAGACACAGCTGTGCTTCCGCGAGGTCTTGG	1680
Db	1621	TGTGTGGATGATGATGTGTCCCGGAGCGAAGACACAGCTGTGCTTCCGCGAGGTCTTGG	1680
QY	1681	TTTCTGGCATATGGGGCCAGAGCCGGTGTGTGTGTCTGTCTTCTGTCTGTAGCTGGCGTGG	1740
Db	1681	TTTCTGGCATATGGGGCCAGAGCCGGTGTGTGTGTGTCTGTCTTCTGTCTGTAGCTGGCGTGG	1740

QY	1741	GGCTTGTGCTGCGTCTGTTGGGGCTGTTCCTTCAACCAATCGGGAACGCCACTGTTAG	1800
Db	1741	GGCTTGTGCTGCGTCTGTTGGGGCTGTTCCTTCAACCAATCGGGAACGCCACTGTTAG	1800
QY	1801	GCCTCGGGGGGGGGCCCTGAGGCTGCTTGGCCCTGATGCTGCTGGGCGTGGCTGCTCAGC	1860
Db	1801	GCCTCGGGGGGGGGCCCTGAGGCTGCTTGGCCCTGATGCTGCTGGGCGTGGCTGCTCAGC	1860
QY	1861	GTCCCTCTGTCCCTTGCCAGCCACGCCCTGCCCCGATGCTTGCCAGCCAGCCCTTGTCC	1920
Db	1861	GTCCCTCTGTCCCTTGCCAGCCACGCCCTGCCCCGATGCTTGCCAGCCAGCCCTTGTCC	1920
QY	1921	CACCTCCCGCTCACGGGCTCTGAGCACACTTCTCTCAGAGCGCCGAGATCTTGCTG	1980
Db	1921	CACCTCCCGCTCACGGGCTCTGAGCACACTTCTCTCAGAGCGCCGAGATCTTGCTG	1980
QY	1981	GAGTCAGAACTGCGCTTGAGCTGTGGGAGAGCCGGCTGATGAGCGGCGTGGGGGGCCCTGG	2040
Db	1981	GAGTCAGAACTGCGCTTGAGCTGTGGGAGAGCCGGCTGATGAGCGGCGTGGGGGGCCCTGG	2040
QY	2041	GCCTGCGTGGTGTGCTGCTGAGCCATGCTGTTGGAGTGTGCACACTGTGCACCTGATCTG	2100
Db	2041	GCCTGCGTGGTGTGCTGCTGAGCCATGCTGTTGGAGTGTGCACACTGTGCACCTGATCTG	2100
QY	2101	GTGGCTTCCCGCCGGAGGTGTGACGGACTGCGACATGCTCTCCACAGGAGGCGCTGTG	2160
Db	2101	GTGGCTTCCCGCCGGAGGTGTGACGGACTGCGACATGCTCTCCACAGGAGGCGCTGTG	2160
QY	2161	CACATCGGCACACGCTCTCTGGGTACGTTTGGGCTTATGCGCAGCCAAATGCCACGCTG	2220
Db	2161	CACATCGGCACACGCTCTCTGGGTACGTTTGGGCTTATGCGCAGCCAAATGCCACGCTG	2220
QY	2221	GCCTTCTCTGCTCTCTGAGGCACTTTCCTGTCGAGCCAGCCGGGCTGTACAAACCT	2280
Db	2221	GCCTTCTCTGCTCTCTGAGGCACTTTCCTGTCGAGCCAGCCGGGCTGTACAAACCT	2280
QY	2281	GCCCGTGCCCTCACCTTGGCATGTGCGCTACTCATCACTCGGGTCTCTTTGTGCC	2340
Db	2281	GCCCGTGCCCTCACCTTGGCATGTGCGCTACTCATCACTCGGGTCTCTTTGTGCC	2340
QY	2341	CTCCGGGCATGTGACAGGTGTCTCAGGCCCGGCTGACATGAGGGGCGCTCTGTCTC	2400
Db	2341	CTCCGGGCATGTGACAGGTGTCTCAGGCCCGGCTGACATGAGGGGCGCTCTGTCTC	2400
QY	2401	TGTGTCTGAGGCACTGCGGTGCTTCCACTGCCAGGTGTACTCTGTATGTGGCAG	2460
Db	2401	TGTGTCTGAGGCACTGCGGTGCTTCCACTGCCAGGTGTACTCTGTATGTGGCAG	2460
QY	2461	CCAGGGCTCAACACACCCCGAGTCTTCTGAGGGGGCCCTGSGGATGCCAAGGCCAG	2520
Db	2461	CCAGGGCTCAACACACCCCGAGTCTTCTGAGGGGGCCCTGSGGATGCCAAGGCCAG	2520
QY	2521	AATGACGGGAACACAGGAATCAGGGGAAACATGATGA	2559
Db	2521	AATGACGGGAACACAGGAATCAGGGGAAACATGATGA	2559
RESULT 4			
US-09-897-427A-7			
; Sequence 7, Application US/09897427A			
; Patent No. US20020160424n1			
; GENERAL INFORMATION:			
; APPLICANT: ADLER, JON ELLIOT			
; APPLICANT: LI, XIADONG			
; APPLICANT: STRZEMSKI, LENA			
; APPLICANT: XU, HONG			
; APPLICANT: EHEVERRI, FERNANDO			
; TITLE OF INVENTION: T1R HETERO-OLIGOMERIC TASTE RECEPTORS			
; FILE REFERENCE: 078003-0282558			
; CURRENT APPLICATION NUMBER: US/09/897,427A			
; CURRENT FILING DATE: 2001-07-03			
; NUMBER OF SEQ ID NOS: 10			
; SOFTWARE: PatentIn Ver. 2.1			

```

; SEQ ID NO 7
; LENGTH: 2559
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-897-427A-7

```

Query Match	99.9%	Score 2557.4	DB 9	Length 2559
Best Local Similarity	100.0%	Pred.No. 0		
Matches 2558	Conservative 0	Mismatches 1	Indels 0	Gaps 0

OY	1	ATGCGAGGACCCGCGTCTGTCGAGGAGGCTACGCTTCGAGGCTGTCGACACCCGAGGACGGGG	60
Db	1	ATGCTGGGACCCCTGCTGTCTCTGGGCTCAAGCCTCTGGGCTCTCTGCAACCCGAGGACGGGG	60
OY	61	GCCCCATTTGTCGCTGTACACAGCAACTTAAAGATGAAGGGGGAGTACGTCTGGGGGGCTG	120
Db	61	GCCCCATTTGTCGCTGTACACAGCAACTTAAAGATGAAGGGGGAGTACGTCTGGGGGGCTG	120
OY	121	TTCCCCCTGGGGCGAGGGCCGAGAGAGGCTTGGCTCCGGCAGCCGGAGACAGGGCCAGAGCCCT	180
Db	121	TTCCCCCTGGGGCGAGGGCCGAGAGAGGCTTGGCTCCGGCAGCCGGAGACAGGGCCAGAGCCCT	180
OY	181	GTGTGCACACAGGTTCTTCCTCAACGGGCTGCTCTGGGACACTGGGCGATGAAATAATGGCCGTG	240
Db	181	GTGTGCACACAGGTTCTTCCTCAACGGGCTGCTCTGGGACACTGGGCGATGAAATAATGGCCGTG	240
OY	241	GAGAGATCAACAAACAACTGGAATCTGCTGCGCCGGGCTGCGCCCTACGACCTCTTTT	300
Db	241	GAGAGATCAACAAACAACTGGAATCTGCTGCGCCGGGCTGCGCCCTACGACCTCTTTT	300
OY	301	GATTAGTGTCTGGAGGCTGTGTGTGGCCATTGAAGCCCAAGCTCATGTTCTGTGGCCAAAGCA	360
Db	301	GATTAGTGTCTGGAGGCTGTGTGTGGCCATTGAAGCCCAAGCTCATGTTCTGTGGCCAAAGCA	360
OY	361	GGCAGCCCGCAGATGGCGGCTACTGTCGCAATACAGCAGATACAGCAGCCCGGTCGTGGCT	420
Db	361	GGCAGCCCGCAGATGGCGGCTACTGTCGCAATACAGCAGATACAGCAGCCCGGTCGTGGCT	420
OY	421	GTGATCGGGGCCCACTGTCGAGAGGCTGCGCCATGGTCAACCGGCAAGTTCTTCAGCTTCTTC	480
Db	421	GTGATCGGGGCCCACTGTCGAGAGGCTGCGCCATGGTCAACCGGCAAGTTCTTCAGCTTCTTC	480
OY	481	CTCATGCCCCCAGGTGAGTACGATGAGGTGTACATGGAAGTGTCTGAGCGCCCGGAGACCTTC	540
Db	481	CTCATGCCCCCAGGTGAGTACGAGGTGTACATGGAAGTGTCTGAGCGCCCGGAGACCTTC	540
OY	541	CCCTCTCTTCGCGAGCGGTGGCCCAAGGCAAGCTGTGACAGCTGAGAGGGCCGGCGGAGACTG	600
Db	541	CCCTCTCTCTTCGCGAGCGGTGGCCCAAGGCAAGCTGTGAGAGGGCCGGCGGAGACTG	600
OY	601	CTGCAAGAGTTCTGGGTGGAACGTGGGTGGCCCTCTGGGCAAGCAGCAGAGTATGAGGCGGG	660
Db	601	CTGCAAGAGTTCTGGGTGGAACGTGGGTGGCCCTCTGGGCAAGCAGCAGAGTATGAGGCGGG	660
OY	661	CAGGGCTTGAGACATCTTCTGCGGCTCTGGCCGCGGACAGCGGCACTCATCTGCGCACAGAG	720
Db	661	CAGGGCTTGAGACATCTTCTGCGGCTCTGGCCGCGGACAGCGGCACTCATCTGCGCACAGAG	720
OY	721	GGCCGTGGGCGCGTGGCCCGGTGGCGGATCACTCGGCGGCTGGGGAAGGTGACAGAGCTCTG	780
Db	721	GGCCGTGGGCGCGTGGCCCGGTGGCGGATCACTCGGCGGCTGGGGAAGGTGACAGAGCTCTG	780
OY	781	CACCAAGTGAACACAGACAGCGTGCAGGTGTGCTGCTTCGCGCTCGGTGACAGCGCGC	840
Db	781	CACCAAGTGAACACAGACAGCGTGCAGGTGTGCTGCTTCGCGCTCGGTGACAGCGCGC	840
OY	841	CACGCGCTTCTCACTACAGCATCAGCAGCAGGCTCTCGCCCAAGGTGTGGGTGGCCAGC	900
Db	841	CACGCGCTTCTCACTACAGCATCAGCAGCAGGCTCTCGCCCAAGGTGTGGGTGGCCAGC	900
OY	901	GAGGCTCTGGGTGACTCTGACACTGTATGGGGGTGCCCGGCAATGGCCCAATGGGACG	960
Db	901	GAGGCTCTGGGTGACTCTGACACTGTATGGGGGTGCCCGGCAATGGGACG	960

[illegible]

Db 1230 GCCAGGCGCTGACAAACCTTTCAGTGAACGCTCAGGCTGCCCGCAGAGACCC 1289
QY 1261 GTGAAGCCCTGGACAGTCTGGAGAACATGTACAACTGACCTTCCAGTGGGGGAGCTG 1320
Db 1290 GTGAAGCCCTGGACAGTCTGGAGAACATGTACAACTGACCTTCCAGTGGGGGAGCTG 1349
QY 1321 CCCTGCGGTTTCAGACAGCAGCGAAGATGACATGAGTACGACCTGAGCTGGGCTG 1380
Db 1350 CCCTGCGGTTTCAGACAGCAGCGAAGATGACATGAGTACGACCTGAGCTGGGCTG 1409
QY 1381 TGGAGGCGCTAGTGGCCAGGCTCCAGACGCTGGGCAAGTTCACAGGAGGCTCAGAGCA 1440
Db 1410 TGGAGGCGCTAGTGGCCAGGCTCCAGACGCTGGGCAAGTTCACAGGAGGCTCAGAGCA 1469
QY 1441 GAGCGCCCTGAAGATCCGCTGGACACGCTGTACAAACAGAACCCGCTGCGGCTGCTG 1500
Db 1470 GAGCGCCCTGAAGATCCGCTGGACACGCTGTACAAACAGAACCCGCTGCGGCTGCTG 1529
QY 1501 CGGCACTGTCAGAGAGGCGCAGGTGCGCGGCTCAAGGGGTTCCACTCTCTGCTAGAC 1560
Db 1530 CGGCACTGTCAGAGAGGCGCAGGTGCGCGGCTCAAGGGGTTCCACTCTCTGCTAGAC 1589
QY 1561 TGTGTGAGCTGAGAGGCGGAGCTACCGGAAACCCAGACGACATGCGCTGACCTT 1620
Db 1590 TGTGTGAGCTGAGAGGCGGAGCTACCGGAAACCCAGACGACATGCGCTGACCTT 1649
QY 1621 TGTGTGAGCTGAGAGTGTGCTCCGAGAGAGACACAGCTCTTCCGCGGAGCTGCTG 1680
Db 1650 TGTGTGAGCTGAGAGTGTGCTCCGAGAGAGACACAGCTCTTCCGCGGAGCTGCTG 1709
QY 1681 TTCTGTGAGTGGGAGGAGCGGCTGTGCTGTCTGCTGTCTGCTGTGCTGTGCTGTG 1740
Db 1710 TTCTGTGAGTGGGAGGAGCGGCTGTGCTGTCTGCTGTCTGCTGTGCTGTGCTGTG 1769
QY 1741 GGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 1800
Db 1770 GGCTTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 1829
QY 1801 GGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1860
Db 1830 GGCTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1889
QY 1861 GTCTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1920
Db 1890 GTCTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1949
QY 1921 CACCTCCGCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1980
Db 1950 CACCTCCGCTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2009
QY 1981 GAGTCAGAACTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2040
Db 2010 GAGTCAGAACTGCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 2069
QY 2041 GCCGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 2100
Db 2070 GCCGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 2129
QY 2101 GTGAGCTTCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2160
Db 2130 GTGAGCTTCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2189
QY 2161 CACTGCGGACAGAGCTCTGGGTGAGCTTGGCTAGGAGGAGGAGGAGGAGGAGGAGGAGG 2220
Db 2190 CACTGCGGACAGAGCTCTGGGTGAGCTTGGCTAGGAGGAGGAGGAGGAGGAGGAGGAGG 2249
QY 2221 GCGTTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 2280
Db 2250 GCGTTTCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTG 2309
QY 2281 GCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2340

Db 2310 GCCGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2369
QY 2341 CTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2400
Db 2370 CTCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2429
QY 2401 TGTGTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2460
Db 2430 TGTGTCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2489
QY 2461 CCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2520
Db 2490 CCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2549
QY 2521 AATGACGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2559
Db 2550 AATGACGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2588

RESULT 6
US-10-282-837-15
; Sequence 15, Application US/10282837
; Publication No. US20030082738A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; TITLE OF INVENTION: 1983, 52881, 2398, 45449, 50289, AND
; FILE REFERENCE: 10448-020001
; CURRENT APPLICATION NUMBER: US/10/282,837
; CURRENT FILING DATE: 2002-10-29
; PRIOR APPLICATION NUMBER: US/09/796,338
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,059
; PRIOR FILING DATE: 2000-02-29
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 15
; LENGTH: 2559
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-282-837-15

Query Match 99.9%; Score 2555.8; DB 9; Length 2559;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGCTGGGCGCTGTGCTGTGAGGCTGAGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
Db 1 ATGCTGGGCGCTGTGCTGTGAGGCTGAGGCTGAGGCTGAGGCTGCTGCTGCTGCTGCTGCTGCTG 60
QY 61 GCCCATTTGTGCTGTGCTGACAGCACTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
Db 61 GCCCATTTGTGCTGTGCTGACAGCACTTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 120
QY 121 TTCCCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
Db 121 TTCCCTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
QY 181 GTGTGACAGAGGCTGTCTCTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
Db 181 GTGTGACAGAGGCTGTCTCTCAAAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 240
QY 241 GAGGAGATCAACAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
Db 241 GAGGAGATCAACAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 300
QY 301 GATAGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
Db 301 GATAGGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 360
QY 361 GGCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420
Db 361 GGCAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 420

421 GTCATCGGGCCCACTCGTCAGAGCTGCCATGTCACCGGCAAGTTCTTCAGCTTCTTC 480
421 GTGATCGGGCCCACTCGTCAGAGCTGCCATGTCACCGGCAAGTTCTTCAGCTTCTTC 480
481 CTGATGCCCAAGTACAGTACGAGTGTGATGAGAGTGTGAGAGTGTGAGAGTGTGAG 540
481 CTGATGCCCAAGTACAGTACGAGTGTGATGAGAGTGTGAGAGTGTGAGAGTGTGAG 540
541 CCTCTCTTCTTCGCGACCTGCGCAGCGACCTGTGCACTACAGGCTCCCGCGAGCTG 600
541 CCTCTCTTCTTCGCGACCTGCGCAGCGACCTGTGCACTACAGGCTCCCGCGAGCTG 600
601 CTGCAAGAGTTGGGCTGGAGTGGGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 660
601 CTGCAAGAGTTGGGCTGGAGTGGGAGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGG 660
661 CAGGAGCTTGAGCATCTTCTCGGCGCTGGCGGCGGAGCGGAGCATCTGATCGCGACGAG 720
661 CAGGAGCTTGAGCATCTTCTCGGCGCTGGCGGCGGAGCGGAGCATCTGATCGCGACGAG 720
721 GGCCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
721 GGCCTGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 780
781 CACCAAGTGAACAGAGCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 840
781 CACCAAGTGAACAGAGCAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 840
841 CAGCGCTCTTCTTCACTACAGCATCAGCAGAGGCTCTCGCCCAAGGTTGGGTGGCCAGC 900
841 CAGCGCTCTTCTTCACTACAGCATCAGCAGAGGCTCTCGCCCAAGGTTGGGTGGCCAGC 900
901 GAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
901 GAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 960
961 GTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
961 GTGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1020
1021 CACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
1021 CACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1080
1081 CTGAGAGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1140
1081 CTGAGAGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1140
1141 GTGAGAGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1200
1141 GTGAGAGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1200
1201 GCGCAGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1260
1201 GCGCAGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1260
1261 GTGAGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1320
1261 GTGAGAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1320
1321 CCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
1321 CCGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1380
1381 TGGCAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1440
1381 TGGCAGAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 1440
1441 GAGCGCTGAAGATCGCTGCGTGGCAGACAGCTGTGACAAACAGAGAGGCGCTGTGCTG 1500
1441 GAGCGCTGAAGATCGCTGCGTGGCAGACAGCTGTGACAAACAGAGAGGCGCTGTGCTG 1500

1501 CCGCAGTGTGAG 1560
1501 CCGCAGTGTGAG 1560
1561 TGTGTGAGTGTGAG 1620
1561 TGTGTGAGTGTGAG 1620
1621 TGTGTGAGTGTGAG 1680
1621 TGTGTGAGTGTGAG 1680
1681 TTCTGTGAGTGTGAG 1740
1681 TTCTGTGAGTGTGAG 1740
1741 GGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
1741 GGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1800
1801 GGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
1801 GGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1860
1861 GTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
1861 GTCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1920
1921 CACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
1921 CACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1980
1981 GAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2040
1981 GAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2040
1981 GAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2040
2041 GGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
2041 GGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2100
2101 GTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2160
2101 GTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGT 2160
2161 CACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2220
2161 CACTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2220
2221 GGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280
2221 GGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280
2281 GGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
2281 GGCCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2340
2341 GTCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2400
2341 GTCTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2400
2401 TGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2460
2401 TGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2460
2461 CAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2520
2461 CAGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2520
2521 AATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2559
2521 AATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAG 2559

Db	1801	GCCTCGGGGGGGGCCCTGGGCTCTTGGCTGATGTGCTGGGCTGTCTGCTCAAC	1860
Qy	1861	GTCTCTCTGTTCCTTGAGCCAGCCAGCCCTGCCGATGCTTGCCAGACGCTTGTCC	1920
Db	1861	GTCCTCTGTTCCTTGCCAGCCAGCCAGCCGCTGAGCTTGCCAGACGCTTGTCC	1920
Qy	1921	CACCTCCGCGTCACGGGCTCCGAGCACACATTCCTGACAGGGGGCCGAGATCTGTG	1980
Db	1921	CACCTCCGCGTCACGGGCTCCGAGCACACATTCCTGACAGGGGGCCGAGATCTGTG	1980
Qy	1981	GAGTCAGAACCTGCTCTGAGCTTGAGCAGACCGGCTGAGTGCTGCTCGGGGGCCCGG	2040
Db	1981	GAGTCAGAACCTGCTCTGAGCTTGAGCAGACCGGCTGAGTGCTGCTCGGGGGCCCGG	2040
Qy	2041	GCCCTGGCTGGTGGTGGCTGGCTGGGCATGCTGGTGGAGAGTGGACCTGTGCACCTGGTACCTG	2100
Db	2041	GCCCTGGCTGGTGGTGGCTGGCTGGGCATGCTGGTGGAGAGTGGACCTGTGTGACCTGGTACCTG	2100
Qy	2101	GTTGGCTCTCCCGCGGAGGGTGGTGAAGAGTGGACATGCTGCCACGAGAGCGCTGGTG	2160
Db	2101	GTTGGCTCTCCCGCGGAGGGTGGTGAAGAGTGGACATGCTGCCACGAGAGCGCTGGTG	2160
Qy	2161	CACCTGCCGACACGCTCTTGAGGTCTTGCGCTTGCGGCTGAGCGACGCAACCAATGCCAGGCTG	2220
Db	2161	CACCTGCCGACACGCTCTTGAGGTCTTGCGCTTGCGGCTGAGCGACGCAACCAATGCCAGGCTG	2220
Qy	2221	GCCCTTCTGAGCTCCGAGGCAATTCCTGAGTGGGAGCGACGCGGGCTGTGTACAAACGT	2280
Db	2221	GCCCTTCTGAGCTCCGAGGCAATTCCTGAGTGGGAGCGACGCGGGCTGTGTACAAACGT	2280
Qy	2281	GCCCGTGGGCTCACTTTGGCATGCTGGGCTACTTACACCTGGGCTTCCCTTTGTGCC	2340
Db	2281	GCCCGTGGGCTCACTTTGGCATGCTGGGCTACTTACACCTGGGCTTCCCTTTGTGCC	2340
Qy	2341	CTCTGGGCAATGTGAGGTGTCTCAAGGCCGCCGCTGACAGATGGGCGCCTCTGCTC	2400
Db	2341	CTCTGGGCAATGTGAGGTGTCTCAAGGCCGCCGCTGACAGATGGGCGCCTCTGCTC	2400
Qy	2401	TGTCTCTGGGCAATGAGGTGCTTCCACCTGAGCCAGGTGTACTGTCATGAGGAG	2460
Db	2401	TGTCTCTGGGCAATGAGGTGCTTCCACCTGAGCCAGGTGTACTGTCATGAGGAG	2460
Qy	2461	CCAGGGCTCAACACCCCGAGTTCCTGAGGAGGGGGCCCTGGGGATGCCCAAGCGAC	2520
Db	2461	CCAGGGCTCAACACCCCGAGTTCCTGAGGAGGGGGCCCTGGGGATGCCCAAGCGAC	2520
Qy	2521	AATGACGGGAACAGAGAAATCAGGGGAAACATGATGA	2559
Db	2521	AATGACGGGAACAGAGAAATCAGGGGAAACATGATGA	2559

RESULT 8
 US-10-282-837-13
 ? Sequence 13, Application US/10282837
 ? Publication No. US20030082738A1
 ? GENERAL INFORMATION:
 ? APPLICANT: Millennium Pharmaceuticals, Inc.
 ? TITLE OF INVENTION: 1983, 52881, 2398, 45449, 50289, AND
 ? TITLE OF INVENTION: 52872, NOVEL G PROTEIN-COUPLED RECEPTORS AND USES THEREFOR
 ? FILE REFERENCE: 10448-020001
 ? CURRENT APPLICATION NUMBER: US/10/282,837
 ? CURRENT FILING DATE: 2002-10-29
 ? PRIOR APPLICATION NUMBER: US/09/796,338
 ? PRIOR FILING DATE: 2001-02-28
 ? PRIOR APPLICATION NUMBER: US 60/186,059
 ? PRIOR FILING DATE: 2000-02-29
 ? NUMBER OF SEQ ID NOS: 26
 ? SOFTWARE: FastSeq for Windows Version 4.0
 ? SEQ ID NO 13
 ? LENGTH: 3489
 ? TYPE: DNA
 ? ORGANISM: Homo sapiens

```

: FEATURE:
: NAME/KEY: CDS
: LOCATION: (52)...(2607)
:
: FEATURE:
: NAME/KEY: misc_feature
: LOCATION: (1)...(3489)
: OTHER INFORMATION: n = A,T,C or G
US-10-282-837-13

Query Match          99.9%   Score 2555.8;   DB 9;   Length 3489;
Best Local Similarity 99.9%   Pred. No. 0;
Matches 2557;   Conservative 0;   Mismatches 2;   Indels 0;   Gaps 0;

```

Query Match		99.99%	Score 2555.8	DB 9	Length 3489
Best Local Similarity		99.99%	Pred. No. 0		
Matches 2557		Conservative	0	Mismatches	2
				Indels	0
				Gaps	0
QY	1	ATGCTGGCCCTGCTGTCTCTGTCGGCCCTGACGCTCTGGGCTCTCTCTGACACCTGGGACGGG	60		
Db	52	ATGCTGGCCCTGCTGTCTCTGTCGGCCCTGACGCTCTGGGCTCTCTCTGACACCTGGGACGGG	1111		
QY	61	GGCCCATTTGTCCTGTCACACCACTTATGATGAAGGGGGGACTACGTGCGGGGGGCTG	120		
Db	112	GGCCCATTTGTCCTGTCACACCACTTATGATGAAGGGGGGACTACGTGCGGGGGGCTG	1717		
QY	121	TTCCCCCTGGGCGGAGGGCCGAGAGAGGTGGCCCTCCGACCGCGACACGGCCACAGAGCTT	180		
Db	172	TTCCCCCTGGGCGGAGGGCCGAGAGAGGTGGCCCTCCGACCGCGACACGGCCACAGAGCTT	2313		
QY	181	GTTGTCACACAGGTTCTCTCTCAACGGCCCTGCTCTGGGCTACTGGCCATAGAAATATGGCCGTG	240		
Db	232	GTTGTCACACAGGTTCTCTCTCAACGGCCCTGCTCTGGGCTACTGGCCATAGAAATATGGCCGTG	2919		
QY	241	GAGGAGATCAACCAACAACTGGATCTGCTGCCCCGGGCTGGGGCTTACGACTCTTTT	3000		
Db	292	GAGGAGATCAACCAACAACTGGATCTGCTGCCCCGGGCTGGGGCTTACGACTCTTTT	3515		
QY	301	GATACGTCTCTGGAGCCCTGCTGGTGGCCATGAAGCCCAAGCCCTCATGTTCTTGGCCCAAGCA	3600		
Db	352	GATACGTCTCTGGAGCCCTGCTGGTGGCCATGAAGCCCAAGCCCTCATGTTCTTGGCCCAAGCA	4111		
QY	361	GGCAGCCCGCAGCATCGCCGCGCTACTGTCACATCAACGAGTACACAGCCCGCTGTGCTGGCT	4200		
Db	412	GGCAGCCCGCAGCATCGCCGCGCTACTGTCACATCAACGAGTACACAGCCCGCTGTGCTGGCT	4717		
QY	421	GTCATCGGGGCCCACTGTCAGAGCTGCGCATGCTCACCGGCAAGTCTTACGTTCTTC	4800		
Db	472	GTCATCGGGGCCCACTGTCAGAGCTGCGCATGCTCACCGGCAAGTCTTACGTTCTTC	5313		
QY	481	CTCATGCCCCCAGTCAGTCTAGTGGGTACATGATGAGTGTGTCGACAGCGCCCGGAGACCTTC	5400		
Db	532	CTCATGCCCCCAGTCAGTCTAGTGGGTACATGATGAGTGTGTCGACAGCGCCCGGAGACCTTC	5919		
QY	541	CCCTCTCTTCTCCACACCGTCCACAGGAGCCGCTGTGACGTGAGCGCGCGCGGAGACTG	6000		
Db	592	CCCTCTCTTCTCCACACCGTCCACAGGAGCCGCTGTGACGTGAGCGCGCGCGGAGACTG	6515		
QY	601	CTGACAGAGATCGGCTGGAACCTGGGTGGCGCCCTGGGGCAGACGACAGATACGGCCGG	6600		
Db	652	CTGACAGAGATCGGCTGGAACCTGGGTGGCGCCCTGGGGCAGACGACAGATACGGCCGG	7111		
QY	661	CAGGGCCCTGACATCTTCTCGGGCCCTGCGCGCGGACAGCGGCACTCTCATCTCGGCACAGAG	7200		
Db	712	CAGGGCCCTGACATCTTCTCGGGCCCTGCGCGCGGACAGCGGCACTCTCATCTCGGCACAGAG	7717		
QY	721	GGCCTGTGGGCGCGTACCCCGGTCGATATCTCGGCGCTGGGGAGAGTGCAGAGAGTCTCTG	7800		
Db	772	GGCCTGTGGGCGCGTACCCCGGTCGATATCTCGGCGCTGGGGAGAGTGCAGAGAGTCTCTG	8313		
QY	781	CACCAAGTGAACCAAGACAGCGTGCAGAGTGGTGTCTGTTGCGCTCCGTGCACAGCCGCC	8400		
Db	832	CACCAAGTGAACCAAGACAGCGTGCAGAGTGGTGTCTGTTGCGCTCCGTGCAGAGCCGCC	8919		
QY	841	CACGCCCTCTTCAACTACAGATACAGACAGGCTCTCGCCCAAGGTGTGGGTGGCCAGC	9000		
Db	892	CACGCCCTCTTCAACTACAGATACAGACAGGCTCTCGCCCAAGGTGTGGGTGGCCAGC	9515		

QY 901 GAGGAGCTGGCTGACCTCTGACCTGATGAGGGGCTCCGGGCTAGGCGGACATGGGACG 960
 |||||
 Db 952 GAGGAGCTGGCTGACCTCTGACCTGATGAGGGGCTCCGGGCTAGGCGGACATGGGACG 1011
 |||||
 QY 961 GTGCTTGGCTTCTCCAGAGGGGCTCCAGCTGACAGATTCCCGCACTAGTGAAGACG 1020
 |||||
 Db 1012 GTGCTTGGCTTCTCCAGAGGGGCTCCAGCTGACAGATTCCCGCACTAGTGAAGACG 1071
 |||||
 QY 1021 CACCTGGGCTTGGGACCGGACCGGCTTCTGCTCTGCTGCTGGGAGAGGAGAGGCT 1080
 |||||
 Db 1072 CACCTGGGCTTGGGACCGGACCGGCTTCTGCTCTGCTGCTGGGAGAGGAGAGGCT 1131
 |||||
 QY 1081 CTGAGAGAGAGAGCTGTGTGGGACAGCGCTGCCGAGTGTGATGATCAGCTGAGAGAC 1140
 |||||
 Db 1132 CTGAGAGAGAGAGCTGTGTGGGACAGCGCTGCCGAGTGTGATGATCAGCTGAGAGAC 1191
 |||||
 QY 1141 GTGAGAGAGAGAGCTTAAATCACCACAGAGCTTCTGCTCTGCTGCTGCTGCTGCTG 1200
 |||||
 Db 1192 GTGAGAGAGAGAGCTTAAATCACCACAGAGCTTCTGCTCTGCTGCTGCTGCTGCTG 1251
 |||||
 QY 1201 GCGGAGGCGCTGCAACAACCTCTTCAAGTGAAGAGCTGAGGCTGCGGCGGAGAGGCGC 1260
 |||||
 Db 1252 GCGGAGGCGCTGCAACAACCTCTTCAAGTGAAGAGCTGAGGCTGCGGCGGAGAGGCGC 1311
 |||||
 QY 1261 GTGAGAGCGCTGAGAGCTTCTGAGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAG 1320
 |||||
 Db 1312 GTGAGAGCGCTGAGAGCTTCTGAGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAG 1371
 |||||
 QY 1321 CCGGCTGCGCTGCAACAAGAGAGAGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAG 1380
 |||||
 Db 1372 CCGGCTGCGCTGCAACAAGAGAGAGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAG 1431
 |||||
 QY 1381 TGGGAGGAGCTGAGAGAGAGAGAGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAG 1440
 |||||
 Db 1432 TGGGAGGAGCTGAGAGAGAGAGAGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAG 1491
 |||||
 QY 1441 GAGGAGCGCTGAGAGAGAGAGAGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAG 1500
 |||||
 Db 1492 GAGGAGCGCTGAGAGAGAGAGAGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAG 1551
 |||||
 QY 1501 GCGGAGTGCAGAGAGAGAGAGAGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAG 1560
 |||||
 Db 1552 GCGGAGTGCAGAGAGAGAGAGAGAGAGAGCTGAGAGAGCTGAGAGAGCTGAGAGAG 1611
 |||||
 QY 1561 TGTGTGAGCTGAGAGAGAGAGAGAGAGTACCGGAGAGAGAGAGAGAGAGAGAGCTTT 1620
 |||||
 Db 1612 TGTGTGAGCTGAGAGAGAGAGAGAGTACCGGAGAGAGAGAGAGAGAGAGAGCTTT 1671
 |||||
 QY 1621 TGTGTGAG 1680
 |||||
 Db 1672 TGTGTGAG 1731
 |||||
 QY 1681 TTTCTGAG 1740
 |||||
 Db 1732 TTTCTGAG 1791
 |||||
 QY 1741 GGCCTTGTGAG 1800
 |||||
 Db 1792 GGCCTTGTGAG 1851
 |||||
 QY 1801 GGCCTTGTGAG 1860
 |||||
 Db 1852 GGCCTTGTGAG 1911
 |||||
 QY 1861 GTGCTTGTGAG 1920
 |||||
 Db 1912 GTGCTTGTGAG 1971
 |||||
 QY 1921 CACCTTCCGCTGAG 1980
 |||||
 Db 1972 CACCTTCCGCTGAG 2031
 |||||
 QY 1981 GAGTACAGAACTGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2040
 |||||

Db 2032 GAGTACAGAACTGCTCTGAG 2091
 |||||
 QY 2041 GCTTGGCTGAG 2100
 |||||
 Db 2092 GCTTGGCTGAG 2151
 |||||
 QY 2101 GTGAGCTTCCCGGAG 2160
 |||||
 Db 2152 GTGAGCTTCCCGGAG 2211
 |||||
 QY 2161 CACTGCGGACAGAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2220
 |||||
 Db 2212 CACTGCGGACAGAGCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2271
 |||||
 QY 2221 GCTTGTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2280
 |||||
 Db 2272 GCTTGTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2331
 |||||
 QY 2281 GCGGAGGCGCTGCAACAACCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2340
 |||||
 Db 2332 GCGGAGGCGCTGCAACAACCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2391
 |||||
 QY 2341 CTCTGAG 2400
 |||||
 Db 2392 CTCTGAG 2451
 |||||
 QY 2401 TGTGTGAG 2460
 |||||
 Db 2452 TGTGTGAG 2511
 |||||
 QY 2461 CCAGAGGCTGCAACAACCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2520
 |||||
 Db 2512 CCAGAGGCTGCAACAACCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2571
 |||||
 QY 2521 AATGAGGAG 2559
 |||||
 Db 2572 AATGAGGAG 2610
 |||||

RESULT 9

US-09-796-338A-13
 ; Sequence 13, Application US/09796338A
 ; Patent No. US20020061522A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals, Inc.
 ; TITLE OF INVENTION: 1983, 52881, 2398, 45449, 50289, AND
 ; FILE REFERENCE: 10448-020001
 ; CURRENT APPLICATION NUMBER: US/09/796,338A
 ; PRIOR FILING DATE: 2001-02-28
 ; PRIOR FILING DATE: 2000-02-29
 ; NUMBER OF SEQ ID NOS: 26
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 13
 ; LENGTH: 3489
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (52)...(2607)
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(3489)
 ; OTHER INFORMATION: n - A,T,C or G
 US-09-796-338A-13

Query Match 99.9%; Score 2555.8; DB 10; Length 3489;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2557; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 1 ATGCTGGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 60
 |||||

Db 52 ATGCTGGCCCTGCTGCTGGGGCTCAGCCCTGCGGCTCTTCCTGACCCCTGGAGCGGG 111
Qy 61 GCCCATTTGTCCTGTCACAGCACTTAGATGAAGGGGAGCTAGCTGTGGGGGGCTG 120
Db 112 GCCCATTTGTCCTGTCACAGCACTTAGATGAAGGGGAGCTAGCTGTGGGGGGCTG 171
Qy 121 TTTCCCTGGGGGAGGGCGGAGAGGCTGAGCTCCGACCGGACAGGCCAGAGCCCT 180
Db 172 TTTCCCTGGGGGAGGGCGGAGAGGCTGAGCTCCGACCGGACAGGCCAGAGCCCT 231
Qy 181 GTGTGACACAGGTTCTCTCTCAAAACGGCTGCTGTGGGACATGAGCATGAATAATGGCCGTG 240
Db 232 GTGTGACACAGGTTCTCTCTCAAAACGGCTGCTGTGGGACATGAGCATGAATAATGGCCGTG 291
Qy 241 GAGAGATCAACAACAAGTCGAGTGTGCTGAGCCGGGGCTGCGGCTGGGCTAGACACTCTTT 300
Db 292 GAGAGATCAACAACAAGTCGAGTGTGCTGAGCCGGGGCTGCGGCTGGGCTAGACACTCTTT 351
Qy 301 GATACGTGCTGGAGCCCTGTGTGGTGGCCATGAAGCCCAAGCCTCATGTTCTGTGGCAAGCA 360
Db 352 GATACGTGCTGGAGCCCTGTGTGGTGGCCATGAAGCCCAAGCCTCATGTTCTGTGGCAAGCA 411
Qy 361 GGCAGCCCGGAGCATGCGCGCTACTGCAACTACAGCAAGTACAGACCCCGTGTGGCT 420
Db 412 GGCAGCCCGGAGCATGCGCGCTACTGCAACTACAGCAAGTACAGACCCCGTGTGGCT 471
Qy 421 GTCATCGGGGCCCACTGCTCAGAGCTCGCCATGTACACGGCAATTTCTTCAGTTCTTTC 480
Db 472 GTCATCGGGGCCCACTGCTCAGAGCTCGCCATGTACACGGCAATTTCTTCAGTTCTTTC 531
Qy 481 CTCATGCCCCAGAGTCAGTACGGTGTAGCATGAGCTGTGAGGGCCCGGGAGACCTTTC 540
Db 532 CTCATGCCCCAGAGTCAGTACGGTGTAGCATGAGCTGTGAGGGCCCGGGAGACCTTTC 591
Qy 541 CCCCTCTTCTCCGACCGGTGCGCCAGCGACGCTGTGACGCTACGCGCCGCGGAGCTG 600
Db 592 CCCCTCTTCTCCGACCGGTGCGCCAGCGACGCTGTGACGCTACGCGCCGCGGAGCTG 651
Qy 601 CTGCAAGAGTTGCGCTGGAACATGAGTGGCCGCTCGGACGAGGAGAGAGTACGGCCG 660
Db 652 CTGCAAGAGTTGCGCTGGAACATGAGTGGCCGCTCGGACGAGGAGAGTACGGCCG 711
Qy 661 CAGGGCTGTAGCATTTCTCGGCCCTGGCCGCGGACGCGGATGTGATCGGACGAG 720
Db 712 CAGGGCTGTAGCATTTCTCGGCCCTGGCCGCGGACGCGGATGTGATCGGACGAG 771
Qy 721 GGCCTGTGGCGGCTGCGCGGTCGATGATCGGGGCTGGGGAAGTGAAGACCTCTG 780
Db 772 GGCCTGTGGCGGCTGCGCGGTCGATGATCGGGGCTGGGGAAGTGAAGACCTCTG 831
Qy 781 CACCAAGTGAACAGAGCAGCTGCAAGGTGTGCTGTTCGCCCTCGTGACCGCCG 840
Db 832 CACCAAGTGAACAGAGCAGCTGCAAGGTGTGCTGTTCGCCCTCGTGACCGCCG 891
Qy 841 CACGCTCTTCTTACATACAGATCAGACAGAGCTGTGCCCCAAGGTGTGGTGGCCAGC 900
Db 892 CACGCTCTTCTTACATACAGATCAGACAGAGCTGTGCCCCAAGGTGTGGTGGCCAGC 951
Qy 901 GAGGCTGTGGCTGACCTGTGACTGTGATGGGGGCTGGCCGAGATGGCCCAATGGGGACG 960
Db 952 GAGGCTGTGGCTGACCTGTGACTGTGATGGGGGCTGGCCGAGATGGCCCAATGGGGACG 1011
Qy 961 GTGCTGTGGCTTCTCAGAGGGGTGCCAGCTGACAGATTTCCCAAGTGTGAAGACG 1020
Db 1012 GTGCTGTGGCTTCTCAGAGGGGTGCCAGCTGACAGATTTCCCAAGTGTGAAGACG 1071
Qy 1021 CACGTGGCCCTGGCCACGACCGGCTTTCCTGCTGCTGGGCGAGAGGAGCAGGCT 1080
Db 1072 CACGTGGCCCTGGCCACGACCGGCTTTCCTGCTGCTGGGCGAGAGGAGCAGGCT 1131
Qy 1081 CTGGAGGAGAGAGCTGGGGGCGAGGCTGGCCGCGAGTGTGATCATCGCTGCAAGAC 1140
Db 1132 CTGGAGGAGAGAGCTGGGGGCGAGGCTGGCCGCGAGTGTGATCATCGCTGCAAGAC 1191

Qy 1141 GTGAGCGAGGGCTAAATCACCACAGAGTTCTGTGTACGACAGTGTATAGCGT 1200
Db 1192 GTGAGCGAGGGCTAAATCACCACAGAGTTCTGTGTACGACAGTGTATAGCGT 1251
Qy 1201 GCCAGGGCCCTGACACAACACTTTCAGTGAACGGCTCAGGTGCCCCGCGAGGACCC 1260
Db 1252 GCCAGGGCCCTGACACAACACTTTCAGTGAACGGCTCAGGTGCCCCGCGAGGACCC 1311
Qy 1261 GTGAAGCCCTGGCACACTCTGTGAGAACTGTAAACCTGACTTTCACAGTGGGGGCTG 1320
Db 1312 GTGAAGCCCTGGCACACTCTGTGAGAACTGTAAACCTGACTTTCACAGTGGGGGCTG 1371
Qy 1321 CCGCTGCGGTTGACAGCAGCGGAACGTGACATGAGATACGACTGAGCTGTGGGT 1380
Db 1372 CCGCTGCGGTTGACAGCAGCGGAACGTGAGATGAGATAGACTGAAGCTGTGGGT 1431
Qy 1381 TGGCAGGGCTCAGTGCCTCAGGCTCCACGACGTGGGAGTTTACGCGACGCTCAGACA 1440
Db 1432 TGGCAGGGCTCAGTGCCTCAGGCTCCACGACGTGGGAGTTTACGCGACGCTCAGACA 1491
Qy 1441 GAGCGCTGAAGATCGGCTGGAGACGTCGTGACAAACGAAACCGGTGTCGGGCTGCG 1500
Db 1492 GAGCGCTGAAGATCGGCTGGAGACGTCGTGACAAACGAAACCGGTGTCGGATGCTG 1551
Qy 1501 CGGCAGTGCAAGAGGGCCAGGTGCGCGGGTCAAGGGGTTCCACTCTGTCTACGAC 1560
Db 1552 CGGCAGTGCAAGAGGGCCAGGTGCGCGGGTCAAGGGGTTCCACTCTGTCTACGAC 1611
Qy 1561 TGTGTGACTGCGAGGGGGGACGTTACGGGCAAAACCCAGACATGCTGACCTTT 1620
Db 1612 TGTGTGACTGCGAGGGGGGACGTTACGGGCAAAACCCAGACATGCTGACCTTT 1671
Qy 1621 TGTGGCCAGATGATGTGTCCCGGAGGAGAACACACAGTGTTCGCCGCGAGTCTGG 1680
Db 1672 TGTGGCCAGATGATGTGTCCCGGAGGAGAACACACAGTGTTCGCCGCGAGTCTGG 1731
Qy 1681 TTTCTGTGAGTGGGGGAGCGGCTGTGCTGTGCTCTCTCTGAGCCCTGCGCTG 1740
Db 1732 TTTCTGTGAGTGGGGGAGCGGCTGTGCTGTGCTCTCTCTGAGCCCTGCGCTG 1791
Qy 1741 GGCCTTGTGTGCTGCTTTTGGGGCTGTTCACATCGGAGACGCCACTGTTCAG 1800
Db 1792 GGCCTTGTGTGCTGCTTTTGGGGCTGTTCACATCGGAGACGCCACTGTTCAG 1851
Qy 1801 GCTCTGGGGGGGCGCCCTGGCCGCTTGGGCTGAGTGGGCTGGGCTGGCTGCGCTAGC 1860
Db 1852 GCTCTGGGGGGGCGCCCTGGCCGCTTGGGCTGAGTGGGCTGGGCTGGCTGCGCTAGC 1911
Qy 1861 GTCTCTGTGTTCCCTGTGGCAGCCAGCCCTGCGGATGCTGGCCAGCAGCCCTTGTTC 1920
Db 1912 GTCTCTGTGTTCCCTGTGGCAGCCAGCCCTGCGGATGCTGGCCAGCAGCCCTTGTTC 1971
Qy 1921 CACCTTCCGCTACAGGGCTGCTGAGCACACTTCTCTGCAAGGGCGGAGATCTTCTG 1980
Db 1972 CACCTTCCGCTACAGGGCTGCTGAGCACACTTCTCTGCAAGGGCGGAGATCTTCTG 2031
Qy 1981 GAGTCAGAACTGCTCTGTGAGTGGGAGACCGGCTGAGTGGGCTGGGCTGGGGGCGC 2040
Db 2032 GAGTCAGAACTGCTCTGTGAGTGGGAGACCGGCTGAGTGGGCTGGGCTGGGGGCGC 2091
Qy 2041 GCTTGGCTGTGTGCTGCTGTGCTGCTGCTGTGTGAGAGTGTGCACTGTGACTG 2100
Db 2092 GCTTGGCTGTGTGCTGCTGTGCTGCTGCTGCTGTGTGAGAGTGTGCACTGTGACTG 2151
Qy 2101 GTGGCTTCCGCGCGAGGTGTGACGACTGTGACATGCTGCCACGAGGCGCTGCTG 2160
Db 2152 GTGGCTTCCGCGCGAGGTGTGACGACTGTGACATGCTGCCACGAGGCGCTGCTG 2211
Qy 2161 CACTGCGGACAGCGCTTCCGGGTGACGCTTCCGGCTTACGCGACGCGCCACTGCGCTG 2220
Db 2212 CACTGCGGACAGCGCTTCCGGGTGACGCTTCCGGCTTACGCGACGCGCCACTGCGCTG 2271

Db 367 CAAACGGCGCTGCTGGGCACTGGCCATGAATAAGGCGGTGAGAGATCAACAACAGT 426
 QY 260 CGATATGCTCCCGGGGCTGGCCGTGAGACCTCTTTGATACGGCTGGAGCCTG 319
 Db 427 CGGATCTGCTCCCGGGGCTGGCCGTGAGACCTCTTTGATACGGCTGGAGCCTG 486
 QY 320 TGTGGCCATGAAGCCAGCCCTCATGTTCTGCGCAAGGAGGAGCCGCGACATCGCG 379
 Db 487 TGTGGCCATGAAGCCAGCCCTCATGTTCTGCGCAAGGAGGAGCCGCGACATCGCG 546
 QY 380 CTTACTGCACTACACGAGTACCAAGCCGCTGTGCTGCTGATGGGCCCCACTGCT 439
 Db 547 CCTACTGCACTACACGAGTACCAAGCCGCTGTGCTGCTGATGGGCCCCACTGCT 606
 QY 440 CAGAGCTGCGCATGTCACCGGCAAGTTCTTCACTTCTTCTCATGCCC----- 489
 Db 607 CAGAGCTGCGCATGTCACCGGCAAGTTCTTCACTTCTTCTCATGCCC----- 666
 QY 490 ----- 489
 Db 667 CCCCCACATGACCCACCCACCCAGCCCTGCCGTGGAGCCCGTGTCAAGAGATG 726
 QY 490 -----CAGTCAAGTACGGTGTCTACATGAGCTGCTGAGCCCGCGAG 535
 Db 727 CTTCTGGCCCTTGACAGTACGCTACGCTGCTACATGAGCTGCTGAGCCCGCGAG 786
 QY 536 CTTTCCCTCTCTTCTGCGCACCGTGGCCAGACCGCTGTGCTGAGCTGAGCCCGCGAG 595
 Db 787 CTTTCCCTCTCTTCTGCGCACCGTGGCCAGACCGCTGTGCTGAGCTGAGCCCGCGAG 846
 QY 596 AGCTGTGCAAGATTTGGCTGGAATGCTGGTGGCCCGCCCTGGGAGCGACAGCAGTACG 655
 Db 847 AGCTGTGCAAGATTTGGCTGGAATGCTGGTGGCCCGCCCTGGGAGCGACAGCAGTACG 906
 QY 656 GCGCGGAGGGCTGAGATTTCTGCGCCCTGGCGGGGAGCGAGGATTCGATCGGCG 715
 Db 907 GCGCGGAGGGCTGAGATTTCTGCGCCCTGGCGGGGAGCGAGGATTCGATCGGCG 966
 QY 716 ACAGAGGCTGTGCTGCGCTGGCCGTGCGATGACTGCGGCTGGGAGAGTGAAGAGC 775
 Db 967 ACAGAGGCTGTGCTGCGCTGGCCGTGCGATGACTGCGGCTGGGAGAGTGAAGAGC 1026
 QY 776 TCTGTCAACAGATTAACAGAGAGCGTGAAGTGTGCTGCTTCCCTCCGTGACG 835
 Db 1027 TCTGTCAACAGATTAACAGAGAGCGTGAAGTGTGCTGCTTCCCTCCGTGACG 1086
 QY 836 CCGGCCAGGCCCTTCTCACTACAGCATACAGACAGGCTTCCGCCAAGGTGTGGTGG 895
 Db 1087 CCGGCCAGGCCCTTCTCACTACAGCATACAGACAGGCTTCCGCCAAGGTGTGGTGG 1146
 QY 896 CCAAGGAGGCTGTGCTGACTGATGCTGCTGATGGGCTGCCGGCATGGCCAGATG 955
 Db 1147 CCAAGGAGGCTGTGCTGACTGATGCTGCTGATGGGCTGCCGGCATGGCCAGATG 1206
 QY 956 GCAAGGCTGTGCTTCTCTCAAGAGGGGTGCCAGCTGACAGAGTTCGCCAGTACGTGA 1015
 Db 1207 GCAAGGCTGTGCTTCTCTCAAGAGGGGTGCCAGCTGACAGAGTTCGCCAGTACGTGA 1266
 QY 1016 AGAGGCTGTGAGAGAGAGAGTGTGGCCAGGCGTCCGCAAGTGTGACTGATCAAGCTGC 1135
 Db 1267 AGAGGCTGTGAGAGAGAGAGTGTGGCCAGGCGTCCGCAAGTGTGACTGATCAAGCTGC 1386
 QY 1327 AGGCTGTGAGAGAGAGAGTGTGGCCAGGCGTCCGCAAGTGTGACTGATCAAGCTGC 1386
 Db 1387 AGAAGCTGTGAGAGAGAGTGTGGCCAGGCGTCCGCAAGTGTGACTGATCAAGCTGC 1446
 QY 1196 GCGTGGGCGAGGCGCTGCAACACCTTCAAGTGAACGCTTACGCGTCCCGCGCAGG 1255

Db 1447 GCGTGGGCGAGGCGCTGCAACACACTTCTAGTGAACGCTTACGCTGCCGCGCAGG 1506
 QY 1256 ACCCGTGAAGCCCTG----- 1271
 Db 1507 ACCCGTGAAGCCCTGCGCAGGTGAGCCCGGAGATGGGGGTGTGCTGCTCTGATGTG 1566
 QY 1272 ----- 1271
 Db 1567 CCCAGGCGACAGGCGCCACGCTGAGCTGGAGGTGGCTGGCGGCTCAGCCCC 1626
 QY 1272 -----GCACTCTGGAAGAACATGTAAACCTGACTTCCAGTGGGCGGCTGC 1321
 Db 1627 TCCCGCGCGCGCAGCTCTGTGGAACATGTAAACCTGACTTCCAGTGGGCGGCTGC 1686
 QY 1322 CGCTGGGTTGAGAGAGAGCGGAACGTGGACATGAGTGAAGCTGAAGCTGAGGTGT 1381
 Db 1687 CGCTGGGTTGAGAGAGAGCGGAACGTGGACATGAGTGAAGCTGAAGCTGAGGTGT 1746
 QY 1382 GGCAGGCGCTAGTCCAGGCTCCACGACGCTGGGCAAGTTCCAGGCGAGCTCAGACAG 1441
 Db 1747 GGCAGGCGCTAGTCCAGGCTCCACGACGCTGGGCAAGTTCCAGGCGAGCTCAGACAG 1806
 QY 1442 AGCGCTGAAGATCCGCTGGGCAAGCTGTACAAAC----- 1476
 Db 1807 AGCGCTGAAGATCCGCTGGGCAAGCTGTGTGACAAACAGTGAAGTGAAGTGGTGTGC 1866
 QY 1477 ----- 1476
 Db 1867 AGGCGTCCCGTGTACCCCGCGGCGAGCGGCGAGCCTGGGGGTGGGCGCTTCCAGT 1926
 QY 1477 -----CAGAGCCCGT 1487
 Db 1927 CTTCCGTGGGATCCAGCGCAGACAGAGCCAGACCCAGCTGTGCGAGAGGCCGT 1986
 QY 1488 GTCCCGGTGGTCCGCGGAGTGCAGAGAGGCGCAGTGGCGCGGCTCAAGGGGTTCACCTC 1547
 Db 1987 GTCCCGGTGGTCCGCGGAGTGCAGAGAGGCGCAGTGGCGCGGCTCAAGGGGTTCACCTC 2046
 QY 1548 CTGCTGTACAGCTGTGTGACCTGCAAGGCGGCGAGCTACCGGCAAAAC----- 1597
 Db 2047 CTGCTGTACAGCTGTGTGACCTGCAAGGCGGCGAGCTACCGGCAAAACAGGTGAGCC 2106
 QY 1598 ----- 1597
 Db 2107 GCTTCCCGGAGCGGGGTGGGAACAGACAGGAGGTCTGCAAGTCTGACTC 2166
 QY 1598 -----CAGA 1601
 Db 2167 TGAGACAGAGCCCAAGGCTACAGAGAACACCCAGCGCTTCTCTCTCACAGT 2226
 QY 1602 CGACATGCGCTGACCTTTTGTGGCCAGATGATGATGCTCCCGAGGAGCAACAGCCTG 1661
 Db 2227 CGACATGCGCTGACCTTTTGTGGCCAGATGATGATGCTCCCGAGGAGCAACAGCCTG 2286
 QY 1662 CTTCCGCGCAGGTGTGCTGCTGCTGATGGGCGAGCCGCTGTGCTGCTGCTGCT 1721
 Db 2287 CTTCCGCGCAGGTGTGCTGCTGCTGATGGGCGAGCCGCTGTGCTGCTGCTGCT 2346
 QY 1722 GCTGCTAGGCTGAGCGCTGGGCGCTTGTGCTGCTGCTTGTGGGCTGTGCTGCTGCT 1781
 Db 2347 GCTGCTAGGCTGAGCGCTGGGCGCTTGTGCTGCTGCTTGTGGGCTGTGCTGCTGCT 2406
 QY 1782 GGAGAGCCCACTGTTAGGCGCTTGGGGGGGCCCTGGCTGTTGGCTGTGCTGCT 1841
 Db 2407 GGAGAGCCCACTGTTAGGCGCTTGGGGGGGCCCTGGCTGTTGGCTGTGCTGCT 2466
 QY 1842 GGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1901
 Db 2467 GGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2526
 QY 1902 GGGCCAGAGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1961
 Db 2527 GGGCCAGAGCGCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2586

QY	1962	GGGGCCGAGATCTTTGGTGGAGTCAAAACAGCTCTGACCTGGGAGAGACCGGCTAGTGG	2021
Db	2587	GGCGGCCGAGATCTTCTGTGAGTACAGAACTGCTTCAGACTCTGGGAGACACCGGCTAGATGG	2646
QY	2022	CTGGCTCTCGGGGGGCGCTCGGGGCGCTGGGCGTGGTGTGCTGGGCGATGCTAGTGGAGTGGC	2081
Db	2647	CTGGCTCTCGGGGGGCGCTCGGGGCGCTGGGCGTGGTGTGCTGGGCGATGCTAGTGGAGTGGC	2706
QY	2082	ACTGTGCACCTGGTACCTGTGGCCCTTCCCGCGAGGTGGTGGAGCGACTGGACATGCT	2141
Db	2707	ACTGTGCACCTGGTACCTGTGGCCCTTCCCGCGAGGTGGTGGAGCGACTGGACATGCT	2766
QY	2142	GCCACAGGAGGCGGTGGTGGACACTGCGGCGACAGGCTCCTGGGTACGCTTGGGCTAGCGCA	2201
Db	2767	GCCACAGGAGGCGGTGGTGGACACTGCGGCGACAGGCTCCTGGGTACGCTTGGGCTAGCGCA	2826
QY	2202	CGCCACCAATGCGCACGCTGGCGCTTTTCCTGCTCTCGGGGACCTTTCCTGGTCCGGAAGCA	2261
Db	2827	CGCCACCAATGCGCACGCTGGCGCTTTTCCTGCTCTCGGGGACCTTTCCTGGTCCGGAAGCA	2886
QY	2262	GCCGGGCTGCTACAAACCGTGCCTCCGCTGGGCTCAACCTTGGCATGCTGGGCTACTTTCATAC	2321
Db	2887	GCCGGGCTGCTACAAACCGTGCCTCCGCTGGGCTCAACCTTGGCATGCTGGGCTACTTTCATAC	2946
QY	2322	CTGGGCTCTCTTTGTGTGCGCCCTCCTGGGCGAATGTGCAAGTGTGCTTCAGGCGCGGCTCA	2381
Db	2947	CTGGGCTCTCTTTGTGTGCGCCCTCCTGGGCGAATGTGCAAGTGTGCTTCAGGCGCGGCTCA	3006
QY	2382	GATGGGCGGCGCTCTGCTGTGTCTCTGGGCGAATCGTGGGTGCTTCCACCTGGCCCCAGGTG	2441
Db	3007	GATGGGCGGCGCTCTGCTGTGTCTCTGGGCGAATCGTGGGTGCTTCCACCTGGCCCCAGGTG	3066
QY	2442	TTACCTGCTCATGCGGCGACGACGAGGCTCAACACCCCGAGTCTTTCCTGGAGGGGGGCC	2501
Db	3067	TTACCTGCTCATGCGGCGACGACGAGGCTCAACACCCCGAGTCTTTCCTGGAGGGGGGCC	3126
QY	2502	TGGGGATGGCCCAAGGCGCAGAAATACGGGAAACACAGGAATTCAGGGGAAACATGATGA	2559
Db	3127	TGGGGATGGCCCAAGGCGCAGAAATACGGGAAACACAGGAATTCAGGGGAAACATGATGA	3184

```

RESULT 12
US-10-261-482-3
? Sequence 3, Application US/10261482
? Publication No. US20030036089A1
? GENERAL INFORMATION:
? APPLICANT: WEI, Ming-Hui et al
? TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
? TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
? TITLE OF INVENTION: PROTEINS, AND USE THEREOF
? FILE REFERENCE: C10008693CN
? CURRENT APPLICATION NUMBER: US/10/261,482
? CURRENT FILING DATE: 2002-10-02
? PRIOR APPLICATION NUMBER: 09/684,393
? PRIOR FILING DATE: 2000-10-10
? PRIOR APPLICATION NUMBER: 60/172,600
? PRIOR FILING DATE: 1999-12-20
? NUMBER OF SEQ ID NOS: 10
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 3
? LENGTH: 8001
? TYPE: DNA
? ORGANISM: Human
US-10-261-482-3
```

Query Match	76.1%;	Score 1947.4;	DB 9;	Length 8001;
Best Local Similarity	82.0%;	Pred. No. 0;		
Matches 2558; Conservative	0;	Mismatches	1;	Indels 560; Gaps 5,

QY 1 ATGCTGGGCGCTGCTGTCTCTGGGCGTCAAGCCTCTGCGCTCTCTGACACCCCTGGGACGGGG 60
|||||
Db 2118 ATGCTGGGCGCTGCTGTCTCTGGGCGTCAAGCCTCTGCGCTCTCTGACACCCCTGGGACGGGG 2177

OY	61	GCCCATTTGTCCTGCTACAGCAACTT	AGATGTAAGAGGGAATAC	TGCTGGGGGGCTG	120
Db	2178	GCCCCATTGTGCTCTGTACAGCAACTT	AGATGTAAGAGGGAATAC	TGCTGGGGGGGGCTG	2237
OY	121	TTCCCCCTGGGGCCGAGGCCGAGAGG	GTGGGCTCCGAGCCGGAACAGGCCCC	AGCAGGCTT	180
Db	2238	TTCCCCCTGGGGCCGAGGCCGAGAGG	CTGGCTCCGAGCCGGAACAGGCCCC	AGCAGGCTT	2297
OY	181	GTGTGCA-----			187
Db	2298	GTGTGCAACACAGGTACAGAGGTGG	AGAGGCTGGGATCAGGGTACACAGGTCTGG		2357
OY	188	-----	-----	-----	199
Db	2358	GTGTCTCTGAGCTGTGGGGCCGAGG	TGGCCATCTGGCTTGTGTGGCCCCAG	GTTCCT	2417
OY	200	CAAAAGGCTGCTCTGTGGGCACTGG	CGCCATGAAATATGCCCTGGAGAGATCA	CAACAAGT	259
Db	2418	CAAAAGGCTGCTCTGTGGGCACTGG	CGCCATGAAATATGCCCTGGAGAGATCA	CAACAAGT	2477
OY	260	CGGATCTGCTGTCCCGGGCTGCGCT	TGGGCTAGACCTCTTTATATACGTCTGG	AGCCTG	319
Db	2478	CGGATCTGCTGTCCCGGGCTGCGCT	TGGGCTAGACCTCTTTATATACGTCTGG	AGCCTG	2537
OY	320	TGTGTGGCATTAAGCCAGGCTCATGT	CTTCCTGGCCAAAGCAGGACGGCAGATCG	CGCG	379
Db	2538	TGTGTGGCATTAAGCCAGGCTCATGT	CTTCCTGGCCAAAGCAGGACGGCAGATCG	CGCG	2597
OY	380	CCCTACTGCACACTACAGCAGTACCA	AGCCCCGTGTGCTGCTGTATCATGGGCC	CACTGCT	439
Db	2598	CCCTACTGCACACTACAGCAGTACCA	AGCCCCGTGTGCTGCTGTATCATGGGCC	CACTGCT	2657
OY	440	CAGAGCTGCGCATGTATCAGCGGCA	AGTCTTCAGCTTCTCTCAATGCC-----		489
Db	2658	CAGAGCTGCGCATGTATCAGCGGCA	AGTCTTCAGCTTCTCTCAATGCC-----		2717
OY	490	-----			489
Db	2718	CCCCCACAATCACCCACCCACCCAG	CCCTGGGGAGCCCTGTGTCAAGAGAT		2777
OY	490	-----	-----	-----	534
Db	2778	GCCCTCTTGGCCCTTTCAGAGTCACT	ACGAGTGTAGTGTGCTGTGAGCCCGGGAG		2837
OY	535	ACCTTCCTCTCTCTTCGACACCGAT	GGCCAGCGACCGTGTACAGTGAACGGCC	CGCG	594
Db	2838	ACCTTCCTCTCTCTTCGACACCGAT	GGCCAGCGACCGTGTGTGAGTGAACGGCC	CGCG	2897
OY	555	GAGCTGCTGCAGAGAGTTGGGCTG	GGAATGGGTGGCCCTTGGGACAGAGAG	AGATAC	654
Db	2898	GAGCTGCTGCAGAGAGTTGGGCTG	GGAATGGGTGGCCCTTGGGACAGAGAG	AGATAC	2957
OY	655	GAGCGGAGAGGACCTGAGACATCTT	CGGGCCCTGGCCGCGGACGCGATCG	CACTCGCG	714
Db	2958	GAGCGGAGAGGACCTGAGACATCTT	CGGGCCCTGGCCGCGGACGCGATCG	CACTCGCG	3017
OY	715	CACGAGGAGCTGTGGCCGCTGGCC	GTGCGATGACTCGCGGCTGGGGAAG	GTGAGAC	774
Db	3018	CACGAGGAGCTGTGTGGCCGCTGGCC	GTGCGATGACTCGCGGCTGGGGAAG	GTGAGAC	3077
OY	775	GTCTCGACACAGGTAACAGAGACAG	AGGTGACAGTGTCTGTGCGCTCGGTG	AG	834
Db	3078	GTCTCGACACAGGTAACAGAGACAG	AGGTGACAGTGTGTGTGTGCTCGGTG	AG	3137
OY	835	GCGCGCCAGCCCTTCAACTACAGAT	CAGCAGGCTCTCGCCCAAGGTGTG	GGGTG	894
Db	3138	GCGCGCCAGCCCTTCAACTACAGAT	CAGCAGGCTCTCGCCCAAGGTGTG	GGGTG	3197
OY	895	GCGACAGAGGCTGTGGCTGACCTGT	AGACATGTGTATGGGGCTGCCGGAT	TGGCCCAATG	954
Db	3198	GCGACAGAGGCTGTGTGACCTGTGT	AGACATGTGTATGGGGCTGCCGGAT	TGGCCCAATG	3257

OY	260	CGGATTCGCTGCCCGGGCTCGGCTGGGGCTAGACACTCTTTTATATACGTCTGGAGCTG	319
Db	571	CGGATTCGCTGCCCGGGCTCGGCTGGGGCTAGACACTCTTTTATACGTCTGGAGCTG	630
OY	320	TGGAGGCATGAACCCAGGCTCATGTCCTGGCCAAAGGACAGGACGCCGACATCGCG	379
Db	651	TGGTGGCCATGAACCCAGGCTCATGTCCTGGCCAAAGGACAGGACGCCGACATCGCG	690
OY	380	CCTACTGCACATACACGACAGTACAGCCCCGTGTGGCTGTCAATCGGGCCCCACTCGT	439
Db	691	CCTACTGCACATACACGACAGTACAGCCCCGTGTGGCTGTCAATCGGGCCCCACTCGT	750
OY	440	CAGAGCTGCCATGTGTACACGGCAAGTTCTTACGCTTCTTCTCATGCC-----	489
Db	751	CAGAGCTGCCATGTGTACACGGCAAGTTCTTACGCTTCTTCTCATGCC-----	810
OY	490	-----	489
Db	811	CCCCCACCATCACCCACCCCAACCAACCCTGCCCGTGGAGCCCCCTTGTGTACGA	870
OY	490	-----	489
Db	871	GAATGCTACATGACACCCACCACCAGCCCTGCCCTGGGAGGCCCTGTGTACAAAGATCTTT	930
OY	490	-----CAGGTGAGCTTACGCTGTAGCATGTGAGCTGTGAGCGCCGGAGACCTTCC	542
Db	931	GGCCTTGACAGGTGAGCTTACGCTGTAGCATGTGAGCTGTGAGCGCCGGAGACCTTCC	990
OY	543	CTCCTTCTTCCGACACCGTGGCCAGCCAGCTGTGACAGCTGACAGCGCCCGGGAGCTCT	602
Db	991	CTCCTTCTTCCGACACCGTGGCCAGCCAGCTGTGACAGCTGACAGCGCCCGGGAGCTCT	1050
OY	603	GCAGAGATTGGGTGGAACACTGGGATGGCCGCCCTGGGACAGACGACGATACGGCCGCA	662
Db	1051	GCAGAGATTGGGTGGAACACTGGGATGGCCGCCCTGGGACAGACGACGATACGGCCGCA	1110
OY	663	GGGCTTGAACATCTTCTCGGCCCTTGCCCGCGGACGCGGCATCTGCATCGCGACGAGG	722
Db	1111	GGGCTTGAACATCTTCTCGGCCCTTGCCCGCGGACGCGGCATCTGCATCGCGACGAGG	1170
OY	723	CGTGTGCGCGCTGCGCCGTCGATGATACATCGCGGCTGGGGAAGTGTGACAGATCTTGTCA	782
Db	1171	CGTGTGCGCGCTGCGCCGTCGATGATACATCGCGGCTGGGGAAGTGTGACAGATCTTGTCA	1230
OY	783	CCAGGTGAACAGAGCAGCTGTGAGGTGTGCTGTTCGCTCCGTCCGTGACAGCGCCCA	842
Db	1231	CCAGGTGAACAGAGCAGCTGTGAGGTGTGCTGTTCGCTCCGTCCGTGACAGCGCCCA	1290
OY	843	CGCCCTTTCACATCAAGCATCAGACAGAGGCTCTTGCCCAAGTGTGGTGGCCACGA	902
Db	1291	CGCCCTTTCACATCAAGCATCAGACAGAGGCTCTTGCCCAAGTGTGGTGGCCACGA	1350
OY	903	GGCCGTGCTGACCTCTTACCTGGTTCATGGGGCTGCCCGGCATGGCCAGATGGGCACGT	962
Db	1351	GGCCGTGCTGACCTCTTACCTGGTTCATGGGGCTGCCCGGCATGGCCAGATGGGCACGT	1410
OY	963	GCTTGGCTTCCCTCAGAGAGGGGTGCCAGCTGCACGAGTTCCTCCAGTACGTGAACGCA	1022
Db	1411	GCTTGGCTTCCCTCAGAGAGGGGTGCCAGCTGCACGAGTTCCTCCAGTACGTGAACGCA	1470
OY	1023	CTTGCCCTTGCGCACCGACCGGCTTGTGCTTGTGCCCTGGGCGAGAGGAGACGGGTCT	1082
Db	1471	CTTGCCCTTGCGCACCGACCGGCTTGTGCTTGTGCCCTGGGCGAGAGGAGACGGGTCT	1530
OY	1083	GGAGGAGGACGTGGTGGGCGACGGCTTGGCCGACGTGTACTGTCACTACAGCTTGCAGACGT	1142
Db	1531	GGAGGAGGACGTGGTGGGCGACGGCTTGGCCGACGTGTACTGTCACTACAGCTTGCAGACGT	1590
OY	1143	GAGCGAGGGCTAAATATACACACAGAGCTTCTGTGCTACGAGAGCTGTATATAGGCTGGC	1202
Db	1591	GAGCGAGGGCTAAATATACACACAGAGCTTCTGTGCTACGAGAGCTGTATATAGGCTGGC	1650
OY	1203	CGAGGCCCTTGACAAACTTTCAGTGTGCAACAGGCTTACAGGCTCCCCCGGCGAGAACCCGT	1262

Db	1651	CCAGGCCCCGCAACACACTCTTCAGTGCACACGGCTCAGGCTCCCGGACAGACCCCGT	1710
QY	1263	GAAGCCCTG-----	1271
Db	1711	GAAGCCCTGCGAAGGTGAGCCCCGGAGATGGGGGTGTGCTGTCTCTGCATGTGCCAGGC	1770
QY	1272	-----	1271
Db	1771	CACCAAGGACAGGCGACCAACGCGTAGCTGGAGGTGGCTGGCGGCTCAGGCCCTCCCGG	1830
QY	1272	---GCAGTCTTGGAGAAACATGTACAACTGACCTTCACGTGGGCGGGCTTCGGCTCG	1328
Db	1831	CCCCAGGCTCTGTGAGAAACATGTACAACTGACCTTCACGTGGGCGGGCTCCCGCTCG	1890
QY	1329	GTTCGACAGCAGGCGAAACGTTGGACATGGAGTACGACCTGAAGCTGTGGGTGTGGCAGG	1388
Db	1891	GTTCGACAGCAGGCGAAACGTTGGACATGGAGTACGACCTGAAGCTGTGGGTGTGGCAGG	1950
QY	1389	CTCAAGTCCCGAGGCTTCACGACAGTGGGGCAGGTTCAACGGCAGCCTCAGACAGAGCGCT	1448
Db	1951	CTCAAGTCCCGAGGCTTCACGACAGTGGGGCAGGTTCAACGGCAGCCTCAGACAGAGCGCT	2010
QY	1449	GAAGATCCGCTGGSCACACGTCGACAAC-----	1476
Db	2011	GAAGATCCGCTGGSCACACGTCGACAACAGTGAAGGTAGAGGTGGGTGTGCCAGGCGTG	2070
QY	1477	-----	1476
Db	2071	CCCGGTAGCCCCCGCGGCGACGGGCGACGCTGGGGGTGGGCGCGTTCACAGTCCCGT	2130
QY	1477	-----CAGAAGCCGCTGCCGG	1494
Db	2131	GGGCATGCCCAAGCGGAGACAGAGCCACAGACCCCAAGCGCTGTGGCGAAGAGCCGCTGCCGG	2190
QY	1495	TGCTTCGGGCGAGTCCAGAGAGGGCCAGAGTCCGCGGCTCAAGGGTTCCACTCTCTGTGC	1554
Db	2191	TGCTTCGGGCGAGTCCAGAGAGGGCCAGAGTCCGCGGCTCAAGGGTTCCACTCTCTGTGC	2250
QY	1555	TACGACTGTGGAGCTCGAGGGGGGAGGTACCGGCAAAAC-----	1597
Db	2251	TACGACTGTGGAGCTCGAGGGGGGAGGTACCGGCAAAACCAAGGTAGCCGCTTCC	2310
QY	1598	-----	1597
Db	2311	CGGCGAGGGGGGGTGGGAGGACGACAGGGGAGGGTCTCTGCAAGTCTGACCTGAGAC	2370
QY	1598	-----CAGAGCAGATC	1608
Db	2371	AGAGCCACAGGATACAAGACGAACACACCGCCCTTCTCTCTCACAGACAGCAATC	2430
QY	1609	GCCTGCACACTTTTGTGGCCAGAGATGATGTGTCCTCGGAGACGAAACACACAGCGCTTCCG	1668
Db	2431	GCCTGCACACTTTTGTGGCCAGAGATGATGTGTCCTCGGAGACGAAACACACAGCGCTTCCG	2490
QY	1669	CGCAGATCTCGGTTCTGTGAGATGGGGGAGACCGGCTGTCTGTGCTGTCTCTCTCTGTG	1728
Db	2491	CGCAGATCTCGGTTCTGTGAGATGGGGGAGACCGGCTGTCTGTGCTGTCTCTCTCTGTG	2550
QY	1729	AGCCGTGCGCTGGGCTTGTGTGCTGTGCTTGTGGGCTGTTCGTTACCAATCGGAGACG	1788
Db	2551	AGCCGTGCGCTGGGCTTGTGTGCTGTGCTTGTGGGCTGTTCGTTACCAATCGGAGACG	2610
QY	1789	CCACGTGTTAGAGGCTGGGGGGGGCCCTTGGCTGTGTTGGGCTGTGTGTGCTTGGGCTTG	1848
Db	2611	CCACGTGTTAGAGGCTGGGGGGGGCCCTTGGCTGTGTTGGGCTGTGTGTGCTTGGGCTTG	2670
QY	1849	GTCTGCTCACAGCTGTCTGTTCCTCGGCGACCCAGCCGTACCGAGTCTTGCGCCAG	1908
Db	2671	GTCTGCTCACAGCTGTCTGTTCCTCGGCGACCCAGCCGTACCGAGTCTTGCGCCAG	2730
QY	1909	CAGCCCTGTGCCACCTCCGCTACGGGCTGCTTGAAGACACTTTCCTGACAGCGGCC	1968


```

Db 2731 CAGCCCTTGTCCACCTCCGCTACGGGCTGCTGAGCACACTCTTCTGACAGCGGCG 2790
QY 1969 GAGATCTTCTGAGTAGAGAACTGCTCTGAGTGGGAGACCGGCTGAGTGGCTGTG 2028
Db 2791 GAGATCTTCTGAGTAGAGAACTGCTCTGAGTGGGAGACCGGCTGAGTGGCTGTG 2850
QY 2029 CGGGGGCCCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2088
Db 2851 CGGGGGCCCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2910
QY 2089 ACCTGTACCTGAGTGGGCTTCCGCGGAGGTGGTGAAGGAGTGGGAGTGGGAGTGGG 2148
Db 2911 ACCTGTACCTGAGTGGGCTTCCGCGGAGGTGGTGAAGGAGTGGGAGTGGGAGTGGG 2970
QY 2149 GAGGCGCTGTGACACTGCGGACACAGCTCTCTGGGTGACGCTTGGGCTTGGGCTTGG 2208
Db 2971 GAGGCGCTGTGACACTGCGGACACAGCTCTCTGGGTGACGCTTGGGCTTGGGCTTGG 3030
QY 2209 AATGCCACGCTGGGCTTCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2268
Db 3031 AATGCCACGCTGGGCTTCTCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3090
QY 2269 TGCTACACACGCTGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2328
Db 3091 TGCTACACACGCTGGGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3150
QY 2329 TCTTTTGTGCCCCCTCTGCGCAATGTGAGTGTGCTCTGAGGCGCGCGCTGAGATGG 2388
Db 3151 TCTTTTGTGCCCCCTCTGCGCAATGTGAGTGTGCTCTGAGGCGCGCGCTGAGATGG 3210
QY 2389 GCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2448
Db 3211 GCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3270
QY 2449 CTCATGGGCGAGCGAGGCTCAACACCCCGAGTTCTTCTGAGGAGGGGCGCTGGGAT 2508
Db 3271 CTCATGGGCGAGCGAGGCTCAACACCCCGAGTTCTTCTGAGGAGGGGCGCTGGGAT 3330
QY 2509 GCCCAAGGCCAAGATGACGGGAGACAGAGAAATCAGGGGAAACATGATGA 2559
Db 3331 GCCCAAGGCCAAGATGACGGGAGACAGAGAAATCAGGGGAAACATGATGA 3381

```

```

RESULT 15
; Sequence 2, Application US/09799629
; Publication No. US20030008344A1
; GENERAL INFORMATION:
; APPLICANT: ADLER, JON ELLIOT
; APPLICANT: ZOZULYA, SERGEY
; APPLICANT: LI, XIADONG
; APPLICANT: O'CONNELL, SHAWN
; APPLICANT: STASZEMSKI, LENA
; TITLE OF INVENTION: T1R TASTE RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: 078003/0277870/RXT
; CURRENT APPLICATION NUMBER: US/09/799,629
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 60/187,546
; PRIOR FILING DATE: 2000-03-07
; PRIOR APPLICATION NUMBER: 60/195,536
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/209,840
; PRIOR FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: 60/214,213
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/226,448
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/259,227
; PRIOR FILING DATE: 2001-01-03
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 2687

```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-799-629-2
Query Match 65.1%; Score 1666; DB 9; Length 2687;
Best Local Similarity 84.7%; Pred. No. 0;
Matches 2070; Conservative 0; Mismatches 0; Indels 374; Gaps 3;
QY 490 CAGGTGACGTACGCTGAGTACATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 549
Db 62 CAGGTGACGTACGCTGAGTACATGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTGAGTGTG 121
QY 550 TTCCGACACGCTGCCAGGACGCTGTGACGTGACGAGGCGCGGAGCTGCTGAGAG 609
Db 122 TTCCGACACGCTGCCAGGACGCTGTGACGTGACGAGGCGCGGAGCTGCTGAGAG 181
QY 610 TTGGGAGGAACTGGGAGGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 669
Db 182 TTGGGAGGAACTGGGAGGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 241
QY 670 AGCATCTTCTCGGCCCTGGGCGGCGGACGCGGATCTGATCGGCAAGAGGAGGAGGAGGAGG 729
Db 242 AGCATCTTCTCGGCCCTGGGCGGCGGACGCGGATCTGATCGGCAAGAGGAGGAGGAGGAGG 301
QY 730 CCGCTGCCCGCTGCCAGTACCTGCGGCTGGGAGAGTGTGACAGCTGCTGACACAGGTG 789
Db 302 CCGCTGCCCGCTGCCAGTACCTGCGGCTGGGAGAGTGTGACAGCTGCTGACACAGGTG 361
QY 790 AATCCAGACGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 849
Db 362 AATCCAGACGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 421
QY 850 TTCACTACACATCAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 909
Db 422 TTCACTACACATCAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 481
QY 910 CTGACCTGTGACCTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 969
Db 482 CTGACCTGTGACCTGTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 541
QY 970 TTCTCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1029
Db 542 TTCTCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 601
QY 1030 CTGGCACAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1089
Db 602 CTGGCACAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 661
QY 1090 GAGGTGTGGGCGCAGGCGCTGCGCAGTGTGATGATGATGATGATGATGATGATGATG 1149
Db 662 GAGGTGTGGGCGCAGGCGCTGCGCAGTGTGATGATGATGATGATGATGATGATGATG 721
QY 1150 GGGCTAATATCAGCAGCAGAGGAGTGTGATGATGATGATGATGATGATGATGATGATG 1209
Db 722 GGGCTAATATCAGCAGCAGAGGAGTGTGATGATGATGATGATGATGATGATGATGATG 781
QY 1210 CTGCAACAACCTTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1269
Db 782 CTGCAACAACCTTCTGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 841
QY 1270 TG----- 1271
Db 842 TGGCAGGTGAGCCCGGAGAGATGGGGGTGTGCTGTCTCTGATGTGCTGAGGAGGAGGAG 901
QY 1272 -----GAG 1275
Db 902 CAGGCGCAACAGGCTGAGCTGAGGAGTGGCTGGGCGCTCAGCCCCCTCCCGCCGAG 961
QY 1276 CTCCTGAGAAATGATCAACCTTCCAGCTGGGCGGAGCTGCGCTGCGCTTGCAG 1335
Db 962 CTCCTGAGAAATGATCAACCTTCCAGCTGGGCGGAGCTGCGCTGCGCTTGCAG 1021
QY 1336 AGCAGCGGAAACGTGACATGAGAGTACGACCTGAAGCTGTGGGTGTGGCAGAGGCTCAGTG 1395

```


THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_A578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:45:20 ; Search time 37.3065 Seconds

(without alignments)
3043.156 Million cell updates/sec

Title: US-09-927-315-15

Perfect score: 4524

Sequence: 1 MGPAYGLSLMALHPQTG.....GPGDAQGQNDGNTGNQKHE 852

Scoring table:

BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 908470 seqs, 133250620 residues

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A.GeneSeq.101002.*
1: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1980.DAT.*
2: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1981.DAT.*
3: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1982.DAT.*
4: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1983.DAT.*
5: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1984.DAT.*
6: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1985.DAT.*
7: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1986.DAT.*
8: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1987.DAT.*
9: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1988.DAT.*
10: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1989.DAT.*
11: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1990.DAT.*
12: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1991.DAT.*
13: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1992.DAT.*
14: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1993.DAT.*
15: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1994.DAT.*
16: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1995.DAT.*
17: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1996.DAT.*
18: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1997.DAT.*
19: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1998.DAT.*
20: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA1999.DAT.*
21: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2000.DAT.*
22: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2001.DAT.*
23: /SIDS2/gcgdata/geneSeq/geneSeq-emb1/AA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4512	99.7	852	22	AA021817
2	4512	99.7	852	23	AA021502
3	4512	99.7	852	23	AAU78587
4	4509	99.7	852	23	AAU73184
5	4500	99.5	850	22	AAE10366
6	4496.5	99.4	863	23	ABR77318
7	3708	82.0	855	23	AAU80493
8	3255	71.9	858	22	AAE10371
9	3229	71.4	858	23	AAU73183
10	3116	68.9	596	23	AAE21818

11	1944	43.0	370	23	AAE21817
12	1769	39.1	1138	23	AA021502
13	1209.5	26.7	840	21	AAV45028
14	1208.5	26.7	840	21	AAV45021
15	1208	26.7	840	21	AAV45026
16	1205.5	26.6	840	21	AAV45027
17	1188.5	26.3	842	22	AAV45022
18	1153	25.5	841	22	AAE10372
19	1153	25.5	841	23	ABR77319
20	1152	25.5	841	23	ABR77319
21	1089.5	24.1	1059	22	AAE11969
22	1082.5	23.6	763	22	AAU00508
23	1068.5	23.6	1078	17	AAE11970
24	1068.5	23.6	1078	19	AAW54846
25	1068.5	23.6	1078	19	AAW38274
26	1068.5	23.6	1078	20	AAV28840
27	1068.5	23.6	1078	20	AAV41780
28	1068.5	23.6	1078	20	AAW89565
29	1068.5	23.6	1078	21	AAV51827
30	1068.5	23.6	1078	21	AAV70325
31	1068.5	23.6	1078	22	AAU02195
32	1068.5	23.6	1078	22	AAV74391
33	1068.5	23.6	1078	23	AAV47822
34	1060	23.4	1085	19	AAW54844
35	1060	23.4	1085	19	AAW38272
36	1060	23.4	1085	20	AAV41778
37	1060	23.4	1085	20	AAW89563
38	1060	23.4	1085	21	AAV51825
39	1057.5	23.4	1088	23	AAV47820
40	1057.5	23.4	1088	17	AAW11888
41	1057.5	23.4	1088	19	AAW54845
42	1057.5	23.4	1088	19	AAW38273
43	1057.5	23.4	1088	20	AAV41779
44	1057.5	23.4	1088	20	AAW89564
45	1057.5	23.4	1088	21	AAV51826

ALIGNMENTS

RESULT 1	AAU08996	standard; Protein; 852 AA.
ID	AAU08996	
AC	AAU08996;	
DT	18-DEC-2001	(first entry)
DE	Human G protein-coupled receptor, GPCR, 50289.	
XX	Human: GPCR; G protein-coupled receptor; 50289; cardiac; antihypertensive; analgesic; cytosolic; antiangiogenic; cardiovascular disorder; pain response disorder; inflammatory disorder; neutral disorder; angina pectoris; myocardial infarction; ischemic heart disease; sudden cardiac death; obesity; hypertensive heart disease; diabetes; prostate cancer-related pain.	
XX	Homo sapiens.	
OS		
XX		
PH	Key	Location/Qualifiers
FT	Peptide	1..20
FT	Domain	1..546
FT	Protein	/label= Extracellular_domain
FT		21..852
FT	Domain	/label= Mature_GPCR50289
FT		61..470
FT		/label= ANF_domain
FT		/note= "Natriuretic peptide (ANF) ligand binding domain"
PN	WO200164882-A2.	

PD 07-SEP-2001.
 XX
 XX 28-FEB-2001; 2001WO-US06543.
 XX
 PR 29-FEB-2000; 2000US-186059P.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Glucksmann MA, Galvin KM, Silos-Santiago I;
 XX
 DR WPI: 2001-589866/66.
 DR N-PSDB: AAS14574.
 XX
 PT Novel G protein coupled receptors and nucleic acids encoding them, for
 PT identifying agents for the treatment of cardiac disorders
 XX
 PS Claim 9; Fig 15; 209pp; English.
 XX
 CC The invention relates to novel human G protein-coupled receptors (GPCR)
 CC named 1983, 52881, 2398, 45449, 50289 and 52872. The novel GPCRs and
 CC nucleic acids encoding them are useful for identifying agents for the
 CC treatment of cardiovascular disorders, angiogenesis-related disorders,
 CC neural disorders, pain response disorders and inflammatory disorders
 CC e.g. atherosclerosis, angina pectoris and myocardial infarction,
 CC ischaemic heart disease, sudden cardiac death, hypertensive heart
 CC disease, diabetes, prostate cancer-related pain, diabetes and obesity.
 CC The present sequence represents GPCR 50289.
 CC
 XX
 SQ Sequence 852 AA;
 Query Match 99.7%; Score 4512; DB 22; Length 852;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 851; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MLGPAVLGLSLMALHPETGAPLCLSQQLRMKGDVILGGLPLGAEAGLRSTRPSP 60
 DB 1 MLGPAVLGLSLMALHPETGAPLCLSQQLRMKGDVILGGLPLGAEAGLRSTRPSP 60
 QY 61 VCTRESSNGLMALAMKMAVEINKKSDLLPGLRGLYDLFTCSBPVYAMRPSLMLAKA 120
 DB 61 VCTRESSNGLMALAMKMAVEINKKSDLLPGLRGLYDLFTCSBPVYAMRPSLMLAKA 120
 QY 121 GSRDIAACNTQYQYPRVLAVIGPHSSSLAMWTGTFSEFPLMPQVSYGASMSLLSARETF 180
 DB 121 GSRDIAACNTQYQYPRVLAVIGPHSSSLAMWTGTFSEFPLMPQVSYGASMSLLSARETF 180
 QY 181 PSFEFTVSDDRVLTAALAEELLOEFEGMNWVALGSDDEYGRGLSTFSLAARGICIAHE 240
 DB 181 PSFEFTVSDDRVLTAALAEELLOEFEGMNWVALGSDDEYGRGLSTFSLAARGICIAHE 240
 QY 241 GLVPLPRADSRGLKVDVYLHOVNOSSVQVYLLFPAVHAHAHLFNYSISSRLSPKVVAS 300
 DB 241 GLVPLPRADSRGLKVDVYLHOVNOSSVQVYLLFPAVHAHAHLFNYSISSRLSPKVVAS 300
 QY 301 EAWLTSDLVGLPGMAQMGTVLGLQRAQLHEPPQYKTHALATDPAFCALGEREOG 360
 DB 301 EAWLTSDLVGLPGMAQMGTVLGLQRAQLHEPPQYKTHALATDPAFCALGEREOG 360
 QY 361 LEEVVGORCPOQCITLQNTSAGLNHQTSVYAAYVSVAOALNHTLOCNAGCPADDP 420
 DB 361 LEEVVGORCPOQCITLQNTSAGLNHQTSVYAAYVSVAOALNHTLOCNAGCPADDP 420
 QY 421 VKPMOLENMTNLTFHVGGLPLRFDSSGNVMEYDLKLMWQSGVPRLDVGRFNGSLRT 480
 DB 421 VKPMOLENMTNLTFHVGGLPLRFDSSGNVMEYDLKLMWQSGVPRLDVGRFNGSLRT 480
 QY 481 ERLKIRMTSNOKPVSRCSHQCEGQYRVKGFHSCCYDCVDEAGSYRONPDDIACF 540
 DB 481 ERLKIRMTSNOKPVSRCSHQCEGQYRVKGFHSCCYDCVDEAGSYRONPDDIACF 540
 QY 541 CGODEMSERSTRCRRRSRLANGEPVALLLLLLSLATGLVLAALGLFVHHRDSPLVQ 600
 DB 541 CGODEMSERSTRCRRRSRLANGEPVALLLLLLSLATGLVLAALGLFVHHRDSPLVQ 600

QY 601 ASGGPLACFGLVCLGVCLSVLPPGQSPARCIAQOPLSHLPLTGCLSTLFLQAAELFV 660
 DB 601 ASGGPLACFGLVCLGVCLSVLPPGQSPARCIAQOPLSHLPLTGCLSTLFLQAAELFV 660
 QY 661 ESELPPLSWADRLSCGLRPMAMLVLLAMLVVALCTWYLVAFPEVYTDHMLPTALV 720
 DB 661 ESELPPLSWADRLSCGLRPMAMLVLLAMLVVALCTWYLVAFPEVYTDHMLPTALV 720
 QY 721 HCTRSMWSFGLAHATNATLAFCLGTFELVRSQGCYNRARGLTPMLAVFTWVSFVP 780
 DB 721 HCTRSMWSFGLAHATNATLAFCLGTFELVRSQGCYNRARGLTPMLAVFTWVSFVP 780
 QY 781 LLANVOVLRPAVOMGALLICVLGILAAFLPRCYLLMRQGLMTPFFLGGPGDAQOQ 840
 DB 781 LLANVOVLRPAVOMGALLICVLGILAAFLPRCYLLMRQGLMTPFFLGGPGDAQOQ 840
 QY 841 NDGNTGNQKHE 852
 DB 841 NDGNTGNQKHE 852
 RESULT 2
 ID AAO21501 standard; Protein; 852 AA.
 AC AAO21501;
 XX
 DT 15-AUG-2002 (first entry)
 XX
 DE Human G-protein coupled receptor (GPCR).
 XX
 KW G-protein coupled receptor; GPCR; gene chip; human; immune response;
 KW chromosomal position; transgenic animal; gene therapy.
 XX
 OS Homo sapiens.
 XX
 FH Key
 FT Region
 FT
 FT Region
 FT
 FT Modified-site
 FT
 FT Region
 FT
 FT Modified-site
 FT
 FT Domain
 FT
 FT Region
 FT
 FT Region
 FT
 FT Region
 FT
 FT Region
 FT
 FT Domain
 FT
 FT Modified-site
 FT
 FT Modified-site
 FT
 FT Modified-site
 FT
 FT Region
 Location/Qualifiers
 20..25
 /note="Myristoylation site"
 69..74
 /note="Myristoylation site"
 85..88
 /note="Asn is N-glycosylated"
 92..97
 /note="Myristoylation site"
 102..105
 /note="Casein kinase II phosphorylation site"
 130..133
 /note="Asn is N-glycosylated"
 137..167
 /label="Helix_1"
 /note="Membrane spanning structure and domain"
 153..155
 /note="Protein kinase C phosphorylation site"
 175..177
 /note="Protein kinase C phosphorylation site"
 175..178
 /note="Casein kinase II phosphorylation site"
 189..191
 /note="Protein kinase C phosphorylation site"
 214..217
 /note="Casein kinase II phosphorylation site"
 234..239
 /note="Myristoylation site"
 256..286
 /label="Helix_3"
 /note="Membrane spanning structure and domain"
 264..267
 /note="Asn is N-glycosylated"
 285..288
 /note="Asn is N-glycosylated"
 289..291


```

|||||
Db 301 EAWLITSDLVKGLPGMAQMGIVLGFLOKGAOLHEPPQYKTHALATDPACFSAALGEREGG 360
QY 361 LEEEDVVGORCPQCCITLQNVASAGLNHHQTFVSVAAYVSAOALHNTLQCNASGCPADP 420
Db 361 LEEEDVVGORCPQCCITLQNVASAGLNHHQTFVSVAAYVSAOALHNTLQCNASGCPADP 420
QY 421 VKPWQLENNYNTLFFHVGGLPLRFDSGGNDMEYDLKLMWQSVPRLDHVGRENGSLRT 480
Db 421 VKPWQLENNYNTLFFHVGGLPLRFDSGGNDMEYDLKLMWQSVPRLDHVGRENGSLRT 480
QY 481 ERLKIRWHTSDNOKPVSRCSQCEGVRRYKGFHSCCYDCVDEAGSYRONPDIDACTF 540
Db 481 ERLKIRWHTSDNOKPVSRCSQCEGVRRYKGFHSCCYDCVDEAGSYRONPDIDACTF 540
QY 541 CGODEWSPERSTRCFRRSRRLANGEPAYVLLLLSLALGLVLAALGLFVHHDSPLVQ 600
Db 541 CGODEWSPERSTRCFRRSRRLANGEPAYVLLLLSLALGLVLAALGLFVHHDSPLVQ 600
QY 601 ASGPPLACFGIVLCGLVCLSVLFFPGQSPARCLAQOPLSHPLTGCLSTFLQAAETIV 660
Db 601 ASGPPLACFGIVLCGLVCLSVLFFPGQSPARCLAQOPLSHPLTGCLSTFLQAAETIV 660
QY 661 ESELPISMADELSCGLNGPMAMLVLLAMLEVALCTWYVAFPEVYTDHMLPTETALV 720
Db 661 ESELPISMADELSCGLNGPMAMLVLLAMLEVALCTWYVAFPEVYTDHMLPTETALV 720
QY 721 HCRTRSWVSEGLAHATNATLAFCLFGLFVRSQPGCYNRAAGLTFMAYFITWSEVP 780
Db 721 HCRTRSWVSEGLAHATNATLAFCLFGLFVRSQPGCYNRAAGLTFMAYFITWSEVP 780
QY 781 LLAVVYVLRPAVOMGALLLCVLTGTLAFLPRCYLLMRQGLMTPPEFLGGEGDAQG 840
Db 781 LLAVVYVLRPAVOMGALLLCVLTGTLAFLPRCYLLMRQGLMTPPEFLGGEGDAQG 840
QY 841 NDGNTGNGKHE 852
Db 841 NDGNTGNGKHE 852

```

RESULT 3
 AAU78587
 ID AAU78587 standard; Protein: 852 AA.
 AC AAU78587;
 XX 18-JUN-2002 (first entry)
 DE Human AXOR79 protein.
 KW AXOR79; human; seven transmembrane; G-protein coupled; receptor;
 KW bacterial; viral; fungal; infection; protozoal; manic depression;
 KW human immunodeficiency virus; HIV; cancer; diabetes; anorexia; bulimia;
 KW Parkinson's disease; heart failure; asthma; allergy; osteoporosis;
 KW neurological disorder; schizophrenia; Huntington's disease.
 XX Homo sapiens.
 OS GB364058-A.
 PN 16-JAN-2002.
 PD 03-MAY-2001; 2001GB-0010887.
 PF 03-MAY-2001; 2000US-0566161.
 PR 05-MAY-2000; 2000US-0566161.
 XX (SMIK) SMITHKLINE BEECHAM CORP.
 PA (SMIK) SMITHKLINE BEECHAM PLC.
 XX Ellis CE, Gattu M;
 PI WPI; 2002-306799/35.
 XX N-PSDB; ABK47345.
 DR

```

XX New G protein-coupled receptor AXOR79, useful for diagnosis and
PT treatment of e.g. infections and cancer -
PS Claim 1; Page 28; 36pp; English.
XX
CC This invention relates to the nucleotide and protein sequences of a
CC novel isolated G-protein coupled 7 transmembrane receptor AXOR79.
CC The DNA and protein sequences of the invention are useful for treating
CC a wide range of diseases such as infections (bacterial, viral, fungal
CC or protozoal, especially human immunodeficiency virus (HIV-1 or -2),
CC pain, cancer, diabetes, anorexia, bulimia, Parkinson's disease,
CC heart failure, asthma, allergy, osteoporosis, neurological disorders
CC (e.g. schizophrenia, manic depression, Huntington's disease) and many
CC others). The sequences are also useful to diagnose, and predict
CC susceptibility to, diseases including detecting mutations or measuring
CC mRNA or protein levels associated with inappropriate activity or levels
CC of AXOR79. The sequences may also be used to screen for agonists or
CC antagonists of AXOR79 and for potential therapeutic agents. The
CC nucleotide sequence can also be used to produce recombinant AXOR79
CC protein which may be used as a vaccine or as source of primers and
CC probes for isolation of related nucleic acids, for chromosomal
CC localisation and in tissue expression studies. The protein sequence
CC can be used to raise specific antibodies which are useful as diagnostic
CC reagents, for isolation/identification of expression clones, for
CC affinity purification of the protein and also for treating the specified
CC diseases. The present sequence represents the human AXOR79 G-protein
CC coupled 7 transmembrane receptor protein of the invention.
XX
SQ Sequence 852 AA;

```

Query Match 99.78; Score 4512; DB 23; Length 852;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 851; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 MLGPVAVIGLSMALHPGTGAPLCLSOOLRKGPYVGLFPGAEAEAGRSTRSSP 60
Db 1 MLGPVAVIGLSMALHPGTGAPLCLSOOLRKGPYVGLFPGAEAEAGRSTRSSP 60
QY 61 VCTRSSNGLLMALAMKMAVEINNKSDLLFGLRGLDYDEPTCSEPVYAMKPSIMPLAKA 120
Db 61 VCTRSSNGLLMALAMKMAVEINNKSDLLFGLRGLDYDEPTCSEPVYAMKPSIMPLAKA 120
QY 121 GSRDIAAYCNTYOYQPRVLAIVGPHSEELAVNTGKFFSFIMPQVSYGASHELLSARETF 180
Db 121 GSRDIAAYCNTYOYQPRVLAIVGPHSEELAVNTGKFFSFIMPQVSYGASHELLSARETF 180
QY 181 PSFRTVPSPDRVOLTAAEELQEEGMMVVAALGSDDEYRGCLSTFSAALAAKICIAHE 240
Db 181 PSFRTVPSPDRVOLTAAEELQEEGMMVVAALGSDDEYRGCLSTFSAALAAKICIAHE 240
QY 241 GLVPLPRADSRRLKQVDVYLHQVQSSVOVYLLFASVHAHAALFNYSISRLSPKRVVVAS 300
Db 241 GLVPLPRADSRRLKQVDVYLHQVQSSVOVYLLFASVHAHAALFNYSISRLSPKRVVVAS 300
QY 301 EAWLITSDLVKGLPGMAQMGIVLGFLOKGAOLHEPPQYKTHALATDPACFSAALGEREGG 360
Db 301 EAWLITSDLVKGLPGMAQMGIVLGFLOKGAOLHEPPQYKTHALATDPACFSAALGEREGG 360
QY 361 LEEEDVVGORCPQCCITLQNVASAGLNHHQTFVSVAAYVSAOALHNTLQCNASGCPADP 420
Db 361 LEEEDVVGORCPQCCITLQNVASAGLNHHQTFVSVAAYVSAOALHNTLQCNASGCPADP 420
QY 421 VKPWQLENNYNTLFFHVGGLPLRFDSGGNDMEYDLKLMWQSVPRLDHVGRENGSLRT 480
Db 421 VKPWQLENNYNTLFFHVGGLPLRFDSGGNDMEYDLKLMWQSVPRLDHVGRENGSLRT 480
QY 481 ERLKIRWHTSDNOKPVSRCSQCEGVRRYKGFHSCCYDCVDEAGSYRONPDIDACTF 540
Db 481 ERLKIRWHTSDNOKPVSRCSQCEGVRRYKGFHSCCYDCVDEAGSYRONPDIDACTF 540
QY 541 CGODEWSPERSTRCFRRSRRLANGEPAYVLLLLSLALGLVLAALGLFVHHDSPLVQ 600

```


Db 541 CGODEMSPERSTRCFRRSRFLAMGEPAYVLLLLLLSLALGLVLAALGLFVHHRDSPLYQ 600
 QY 601 ASGGPLACGLVCLGLVCLSVLLFPQSPACLAQOPLSHPLTGCSTLFLQAAETV 660
 Db 601 ASGGPLACGLVCLGLVCLSVLLFPQSPACLAQOPLSHPLTGCSTLFLQAAETV 660
 QY 661 ESEIPLSNADRLSGCLRGPMAMLVLLAMLVVALCTWTLVAFPEVYTDHMLPTEALV 720
 Db 661 ESEIPLSNADRLSGCLRGPMAMLVLLAMLVVALCTWTLVAFPEVYTDHMLPTEALV 720
 QY 721 HORTSRWSFGLAHATNATLAFCLGTFVLVRSOPGCYNRARGTLFAMLAFTTWSEVP 780
 Db 721 HORTSRWSFGLAHATNATLAFCLGTFVLVRSOPGCYNRARGTLFAMLAFTTWSEVP 780
 QY 781 LLANQVYLPRVAVQMGALLLCVGLTAAHFLPRCTLLMROPLNPEFFLGSGPDAGQ 840
 Db 781 LLANQVYLPRVAVQMGALLLCVGLTAAHFLPRCTLLMROPLNPEFFLGSGPDAGQ 840
 QY 841 NDGNTGNOCKHE 852
 Db 841 NDGNTGNOCKHE 852

RESULT 4
 AAU73184 standard; Protein: 852 AA.

AAU73184;

12-MAR-2002 (first entry)

Human SACL polypeptide.

Human: mouse; SACL: carbohydrate; sweetener; ethanol; alcoholism;
 obesity; diabetes; transgenic embryo; body tissue; body fluid; pancreas;
 blood; tongue; anorectic; antidiabetic; gene therapy;

protein replacement therapy.

Homo sapiens.

Wo200183749-A2.

08-NOV-2001.

25-APR-2001; 2001WO-US13387.

28-APR-2000; 2000US-200794P.

28-JUL-2000; 2000US-22119P.

10-NOV-2000; 2000US-247443P.

(MARIN) WARNER LAMBERT CO.

(MONE-) MONELL CHEM SENSES CENT.

Bechmanov AA, Beauchamp GK, Chatterjee A, De Jong PJ, Li S, Li X;

Omnen JD, Reed DR, Ross D, Tordoff MG;

WPI: 2002-075162/10.

N-PSDB; AAS97395.

Novel isolated polypeptide comprising variant form of mouse or human

SACL polypeptide, and is associated with altered preference for

carbohydrates or other sweeteners, useful for preventing obesity,

diabetes, alcoholism

Claim 7; Page 44-45; 239pp; English.

CC embryos may be used in screening for and identifying agents that induce
 CC or repress function of SACL. Predisposition to diabetes, obesity or
 CC alcoholism can be ascertained by testing any fluid or tissue of a human
 CC (such as blood, pancreas or tongue) for sequence variations of the SACL
 CC gene. A sequence variation of the SACL locus may indicate a
 CC predisposition to diabetes, obesity and/or alcoholism and may provide a
 CC diagnostic mark. The polynucleotide can be detected in a biological
 CC sample by contacting the DNA with a probe to form a hybridisation complex
 CC which is then detected. The sequences represent the human and mouse SACL
 CC polypeptides.

Sequence 852 AA;

Query Match 99.7%; Score 4509; DB 23; Length 852;

Best Local Similarity 99.8%; Pred. No. 0;

Matches 850; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MLGPAYVGLSLMALHPGAPLCLSOOLRMKGDVYLGFLPGBAEAGLSRTRPSP 60
 Db 1 MLGPAYVGLSLMALHPGAPLCLSOOLRMKGDVYLGFLPGBAEAGLSRTRPSP 60
 QY 61 VCTRESSNGLMALAMKAAVEEINKSDLPGLRLGYDLFTDCEPYYAMKPSLMFLAKA 120
 Db 61 VCTRESSNGLMALAMKAAVEEINKSDLPGLRLGYDLFTDCEPYYAMKPSLMFLAKA 120
 QY 121 GSRDIAACNTQYQYPRVIAVIGPHSSBELAMTGFSEFFLMPOVSYGASMLLSARETF 180
 Db 121 GSRDIAACNTQYQYPRVIAVIGPHSSBELAMTGFSEFFLMPOVSYGASMLLSARETF 180
 QY 181 PSFFRTVSDRYOLTAALAEILQEFGMNVAALGSDDEXGROGLSTFSAALARGICIAHE 240
 Db 181 PSFFRTVSDRYOLTAALAEILQEFGMNVAALGSDDEXGROGLSTFSAALARGICIAHE 240
 QY 241 GLVPLPRADDSRLGVQDYDLHOVNOSVQVYLLFRASVHAHALFNYSISSRLSPVYVAA 300
 Db 241 GLVPLPRADDSRLGVQDYDLHOVNOSVQVYLLFRASVHAHALFNYSISSRLSPVYVAA 300
 QY 301 EAMLTSDLYMGLPGMAONGTVLGFQRCAGOLHEFPQYKTHLATDPAFCALGERBOG 360
 Db 301 EAMLTSDLYMGLPGMAONGTVLGFQRCAGOLHEFPQYKTHLATDPAFCALGERBOG 360
 QY 361 LEEDVVGQRCPOQDCITLONVSAGLNHHQTFSVYAAVSVAAQALNHTLOCNASGCPADP 420
 Db 361 LEEDVVGQRCPOQDCITLONVSAGLNHHQTFSVYAAVSVAAQALNHTLOCNASGCPADP 420
 QY 421 VKPQDLLENMYLTFHVGGPLRFDSGQVDMETDKLAWMOGSYPRLDHVGRENGSLRT 480
 Db 421 VKPQDLLENMYLTFHVGGPLRFDSGQVDMETDKLAWMOGSYPRLDHVGRENGSLRT 480
 QY 481 ERLKIRWHTSDNQKPVSRCSROCGOYRVYGFHSCCYDCVDCAGSYRONPDDIACTF 540
 Db 481 ERLKIRWHTSDNQKPVSRCSROCGOYRVYGFHSCCYDCVDCAGSYRONPDDIACTF 540
 QY 541 CGODEMSPERSTRCFRRSRFLAMGEPAYVLLLLLLSLALGLVLAALGLFVHHRDSPLYQ 600
 Db 541 CGODEMSPERSTRCFRRSRFLAMGEPAYVLLLLLLSLALGLVLAALGLFVHHRDSPLYQ 600
 QY 601 ASGGPLACGLVCLGLVCLSVLLFPQSPACLAQOPLSHPLTGCSTLFLQAAETV 660
 Db 601 ASGGPLACGLVCLGLVCLSVLLFPQSPACLAQOPLSHPLTGCSTLFLQAAETV 660
 QY 661 ESEIPLSNADRLSGCLRGPMAMLVLLAMLVVALCTWTLVAFPEVYTDHMLPTEALV 720
 Db 661 ESEIPLSNADRLSGCLRGPMAMLVLLAMLVVALCTWTLVAFPEVYTDHMLPTEALV 720
 QY 721 HORTSRWSFGLAHATNATLAFCLGTFVLVRSOPGCYNRARGTLFAMLAFTTWSEVP 780
 Db 721 HORTSRWSFGLAHATNATLAFCLGTFVLVRSOPGCYNRARGTLFAMLAFTTWSEVP 780
 QY 781 LLANQVYLPRVAVQMGALLLCVGLTAAHFLPRCTLLMROPLNPEFFLGSGPDAGQ 840
 Db 781 LLANQVYLPRVAVQMGALLLCVGLTAAHFLPRCTLLMROPLNPEFFLGSGPDAGQ 840

QY 841 NDGNTGNQKHE 852
 |||||
 Db 841 NDGNTGNQKHE 852

RESULT 5
 AAEL0366
 ID AAEL0366 standard; Protein: 850 AA.
 AC AAEL0366;
 DT 10-DEC-2001 (first entry)
 DE Human taste-cell-specific G protein-coupled receptor, hTIR3 protein.
 XX
 XX Human: taste-cell-specific G protein-coupled receptor; TIR3; drug;
 KM genetic modulation; pharmaceutical; taste sensation; food industry;
 KM chemosensory transduction; chromosome 1p36.2-1p36.33.
 OS Homo sapiens.
 XX
 XX W0200166563-A2.
 PD 13-SEP-2001.
 PF 07-MAR-2001; 2001MO-US07265.
 XX
 XX 07-MAR-2000; 2000US-0187546.
 PR 07-APR-2000; 2000US-0195536.
 PR 06-JUN-2000; 2000US-0209840.
 PR 23-JUN-2000; 2000US-0214213.
 PR 17-AUG-2000; 2000US-0226448.
 PR 03-JAN-2001; 2001US-0259227.
 XX
 XX (SEMO-) SENOMYX INC.
 PA
 XX Adler JE, Zozulya S, O'Connell SM, Li X, Staszewski L;
 PI WPI: 2001-582267/65.
 DR N-PSDB: AAD17509, AAD17517.
 XX
 XX New mammalian taste-cell-specific G protein-coupled receptor
 PT polypeptides for identifying compounds that modulate taste signaling
 PT are useful in food, to modulate the sweet taste of foods or drugs -
 XX
 XX Claim 152; Page 73-74; 119pp; English.

CC The invention relates to mammalian taste-cell-specific G protein-coupled
 CC receptors, TIR and their corresponding cDNA molecules. Taste receptors,
 CC TIR are useful for screening compounds which are used to activate or
 CC modulate chemosensory transduction, such as taste sensation. The
 CC identification and isolation of novel taste receptors and taste
 CC signalling molecules allow for new methods of chemical and genetic
 CC modulation of taste transduction pathways. The taste modulating
 CC compounds are useful in pharmaceuticals and food industries to improve
 CC the taste of a variety of consumer products, or to block undesirable
 CC tastes, e.g., in certain pharmaceuticals. TIRs are also useful in
 CC biochemical assay for identifying tastant (TIR) ligands having binding
 CC specificity for TIR involved in taste signalling. The present sequence is
 CC human taste-cell-specific G protein-coupled receptor, hTIR3 protein.
 CC Human TIR3 gene is localised on chromosome 1p36.2-1p36.33.
 XX
 XX Sequence 850 AA;

Query Match 99.5%; Score 4500; DB 22; Length 850;
 Best Local Similarity 99.6%; Pred. No. 0;
 Matches 84; Conservative 0; Mismatches 1; Indels 2; Gaps 1;

QY 1 MLCGAVLGLSLMALHPGTGAPLCTSOOLRMKGDYVIGLPLGAEPAGRSRTRPSSP 60
 |||||
 Db 1 MLCGAVLGLSLMALHPGTGAPLCTSOOLRMKGDYVIGLPLGAEPAGRSRTRPSSP 60

QY 61 VCTRFSSNGLLMALAMKAAVEINNKSDLLPGLRLGYDLFDTCSEPVVAKPSIMFLAKA 120

Db 61 VCTRFSSNGLLMALAMKAAVEINNKSDLLPGLRLGYDLFDTCSEPVVAKPSIMFLAKA 120
 |||||
 QY 121 GSRDIAAYCNYTOYOPRYLAVICGPHSELAMVYGFSEFLMQVSYGASMELLSARETF 180
 |||||
 Db 121 GSRDIAAYCNYTOYOPRYLAVICGPHSELAMVYGFSEFLMPH--YGASMELLSARETF 178
 |||||
 QY 181 PSFRTVPSPDRVOLTAAEELLOEFGNMVVAALSGDDEYGRQGLSIFSALAANGICIAHE 240
 |||||
 Db 179 PSFRTVPSPDRVOLTAAEELLOEFGNMVVAALSGDDEYGRQGLSIFSALAANGICIAHE 238
 |||||
 QY 241 GLVPLPRADSRIGKVDVYVHOVQSSVQVYLLFASVHAHAALFNYSISRLSPKRWVAS 300
 |||||
 Db 239 GLVPLPRADSRIGKVDVYVHOVQSSVQVYLLFASVHAHAALFNYSISRLSPKRWVAS 298
 |||||
 QY 301 EAMLTSDLVNGLPMAAMGTIVGLFQGAOLHEEPVYKTHLALADPAPCSALGEREG 360
 |||||
 Db 299 EAMLTSDLVNGLPMAAMGTIVGLFQGAOLHEEPVYKTHLALADPAPCSALGEREG 358
 |||||
 QY 361 LEEDVVGQRCPQDCITLQNVASAGLNHQTFSYAAVYVAQALHNTLQCNASGCPAQP 420
 |||||
 Db 359 LEEDVVGQRCPQDCITLQNVASAGLNHQTFSYAAVYVAQALHNTLQCNASGCPAQP 418
 |||||
 QY 421 VKPQQLLENNTLTFHVGGLPLRDSGAVDMETDLKLMWQSSVPLHVGKRNGLRT 480
 |||||
 Db 419 VKPQQLLENNTLTFHVGGLPLRDSGAVDMETDLKLMWQSSVPLHVGKRNGLRT 478
 |||||
 QY 481 ERLKIRHMTSDNOKPVSRGRCOCQEGVRRVKGPHSCCYCQVCEAGSYRONPDIACTE 540
 |||||
 Db 479 ERLKIRHMTSDNOKPVSRGRCOCQEGVRRVKGPHSCCYCQVCEAGSYRONPDIACTE 538
 |||||
 QY 541 CGODEMSPERSTRCFRRRSFELAMGEPAVLLLLLSIALGLVLAALGFVHHRDSEPLVQ 600
 |||||
 Db 539 CGODEMSPERSTRCFRRRSFELAMGEPAVLLLLLSIALGLVLAALGFVHHRDSEPLVQ 598
 |||||
 QY 601 ASGGPLACFGIVCLGLVCLSVLLFPQOPSPARCLAQOPLSHPLTGCLSTFLQAAEIFV 660
 |||||
 Db 599 ASGGPLACFGIVCLGLVCLSVLLFPQOPSPARCLAQOPLSHPLTGCLSTFLQAAEIFV 658
 |||||
 QY 661 ESELPLSMADRLSGCLRGPNAMLVLLAMLEVALCWYVAPPEVVTOMHMLPTFALV 720
 |||||
 Db 659 ESELPLSMADRLSGCLRGPNAMLVLLAMLEVALCWYVAPPEVVTOMHMLPTFALV 718
 |||||
 QY 721 HCRTRSWSPGLAHATNATLAFCLFGLTFVRSOPGCYNRARGLTFMALVFTWSEVP 780
 |||||
 Db 719 HCRTRSWSPGLAHATNATLAFCLFGLTFVRSOPGCYNRARGLTFMALVFTWSEVP 778
 |||||
 QY 781 LLAHVQVVLPRAYOMGALLLCVGLTILAFHLPRCYLLMRPGINTPREFLGGGGRDQOQ 840
 |||||
 Db 779 LLAHVQVVLPRAYOMGALLLCVGLTILAFHLPRCYLLMRPGINTPREFLGGGGRDQOQ 838
 |||||
 QY 841 NDGNTGNQKHE 852
 |||||
 Db 839 NDGNTGNQKHE 850

RESULT 6
 ABB77318
 ID ABB77318 standard; Protein: 863 AA.
 AC ABB77318;
 DT 17-JUN-2002 (first entry)
 DE Human G-protein coupled receptor SRQ ID NO 2.
 XX
 XX Human: G-protein coupled receptor; receptor; anti-HIV; antitumour;
 KM antinflammatory; antiallergic; antianemic; antiasthmatic; virucide;
 KM immunosuppressive; dermatological; nephrotropic; antigout; antithyroid;
 KM cytostatic; neuroproliferative; osteopathic; antiparasitic; antineumatic;
 KM antitachytic; thyromimetic; antidiuretic; ophthalmological; antibacterial;
 KM fungicide; antiparasitic; protozoacide; antihelmintic; antidiabetic;
 KM antitartaroclerotic; hepatotropic; anticonvulsant; anorectic; metabolic;

KW antiemetic; anti diarrhoeic; neuroleptic; cerebroprotective; nootropic;
 KW antiparkinsonian; depilatory; tranquilizer; hypotensive; vasotropic;
 KW cardiant; antianginal; vulnery; proliferative disorder; cancer;
 KW neurological disorder; Alzheimer's disease; Huntington's disease;
 KW Parkinson's disease; multiple sclerosis; meningitis; prion;
 KW cardiovascular disorder; acquired immunodeficiency syndrome; AIDS;
 KW Crohn's disease; diabetes mellitus; rheumatoid arthritis.
 XX Homo sapiens.
 XX WO200198323-A2.
 XX 27-DEC-2001.
 XX 15-JUN-2001; 2001WO-US19354.
 XX 16-JUN-2000; 2000US-212483P.
 XX 23-JUN-2000; 2000US-213950P.
 XX 26-JUN-2000; 2000US-214062P.
 XX 07-JUL-2000; 2000US-216595P.
 XX 14-JUL-2000; 2000US-218936P.
 XX 19-JUL-2000; 2000US-219154P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Lal P, Graul R, Hafalia AJA, Walla NK, Thornton M, Nguyen DB;
 PI Lu Y, Gandhi AR, Patterson C, Kallilick DA, Baughn N, Rammumar J;
 PI Tribouley CM, Lee EA, Ding L, Burford N, Yao MG, Yang J;
 PI Griffin JA;
 XX WPI: 2002-139780/18.
 DR N-PSDB; ABL55952.
 XX Novel G-protein coupled receptor protein and polynucleotides useful for
 PT diagnosing, treating or preventing disorders of cell proliferation e.g.
 PT cancer, neurological and genetic disorder e.g. thalassemia
 XX Claim 47: Page 105-107, 121pp; English.
 XX The invention relates to a G-protein coupled receptor protein (GCREC)
 CC polypeptide (AB877317-AB877326). The GCREC is useful for screening an
 CC agonist/antagonist of GCREC, a compound that specifically binds to GCREC
 CC or that modulates the activity of GCREC. GCREC is also useful as an
 CC immunogen for preparing antibodies which are useful for diagnosing a
 CC condition of disease associated with expression of GCREC in a subject,
 CC for detecting and purifying GCREC from a sample. The GCREC encoding
 CC polynucleotide (ABL55951-ABL55960) is useful for screening for a compound
 CC effective in altering expression of GCREC. GCREC is also useful for
 CC assessing toxicity of a test compound useful for treating a disease or
 CC condition associated with decreased expression or overexpression of
 CC functional GCREC. Examples of disorders include cell proliferative
 CC disorder such as arteriosclerosis, atherosclerosis, hepatitis, mixed
 CC connective tissue disease (MCTD), psoriasis and cancer including
 CC adenocarcinoma, leukemia; a neurological disorder such as epilepsy,
 CC stroke, Alzheimer's disease, Huntington's disease, dementia, Parkinson's
 CC disease, retinitis pigmentosa, multiple sclerosis, bacterial and viral
 CC meningitis, abscess, subdural empyema; prion disease including kuru,
 CC Creutzfeldt-Jakob disease; fatal familial insomnia, neurofibromatosis,
 CC tuberous sclerosis, cerebral palsy, polymyositis; inherited, metabolic,
 CC endocrine, and toxic myopathies; myasthenia gravis, periodic paralysis;
 CC mental disorders including mood, anxiety, and schizophrenic disorders;
 CC seasonal affective disorder (SAD); akathisia, amnesia, cataplexia,
 CC diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses,
 CC Tourette's disorder; cardiovascular disorders such as hypertension,
 CC vasculitis, vascular tumours, congestive heart failure, ischemic heart
 CC disease, myocardial infarction, calcific aortic valve stenosis,
 CC infective endocarditis, endocarditis of systemic lupus erythematosus,
 CC cardiac transplantation; gastrointestinal disorder such as dysphagia,
 CC gastritis, anorexia, nausea, emesis, abdominal angina, infections of
 CC the intestinal tract, peptic ulcer, hepatitis, cirrhosis, diarrhoea,
 CC acquired immunodeficiency syndrome (AIDS) enteropathy, jaundice, Reye's
 CC syndrome, liver infarction; an autoimmune/inflammatory disorder such as
 CC AIDS, Addison's disease, adult respiratory distress syndrome,

CC allergies, amyloidosis, anaemia, asthma, atherosclerosis, autoimmune
 CC thyroiditis, bronchitis, Crohn's disease, diabetes mellitus, thyroiditis,
 CC Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis,
 CC multiple sclerosis, myasthenia gravis, myocardial or pericardial
 CC inflammation, osteoporosis, psoriasis, rheumatoid arthritis,
 CC scleroderma, Sjogren's syndrome, systemic lupus erythematosus, uveitis,
 CC viral, bacterial, fungal, parasitic, protozoal, and helminthic
 CC infections, and trauma; a metabolic disorders such as diabetes,
 CC obesity, and osteoporosis. The GCREC encoding polynucleotide is useful
 CC for gene therapy and for creating knock in humanised animals (pigs) or
 CC transgenic animals (mice or rats) to model human disease.
 XX Sequence 863 AA;
 XX
 XX Query Match 99.4%; Score 4496.5; DB 23; Length 863;
 XX Best Local Similarity 98.6%; Pred. No. 0;
 XX Matches 851; Conservative 0; Mismatches 1; Indels 11; Gaps 1;
 OY 1 MUGPAVLGSLMALHPPTGAPLCLSQQLRMKGDVVLGGPLGAEAGLRSPTRPSP 60
 Db 1 MUGPAVLGSLMALHPPTGAPLCLSQQLRMKGDVVLGGPLGAEAGLRSPTRPSP 60
 OY 61 VCTRESSNGLLMALMKMAVEEINNKSDLPGLRLGYDLFTCSBPVYAMKPSLMEFLAKA 120
 Db 61 VCTRESSNGLLMALMKMAVEEINNKSDLPGLRLGYDLFTCSBPVYAMKPSLMEFLAKA 120
 OY 121 GSRLIATCNCNTQYQPRVLAIVGPHSSSELAWTGTGFSEFLMPQVSYGASMLLSARETF 180
 Db 121 GSRLIATCNCNTQYQPRVLAIVGPHSSSELAWTGTGFSEFLMPQVSYGASMLLSARETF 180
 OY 181 PSFERTVPSDRVQTLAAELLOEFGMNVAALGSODEXRGGLSTFSLAALAGICINHE 240
 Db 181 PSFERTVPSDRVQTLAAELLOEFGMNVAALGSODEXRGGLSTFSLAALAGICINHE 240
 OY 241 GLVPLPRADSRGLKVQVYLHOVNOSSVQVVLLEFSAHAHLNYSISSRLSPRWYAS 300
 Db 241 GLVPLPRADSRGLKVQVYLHOVNOSSVQVVLLEFSAHAHLNYSISSRLSPRWYAS 300
 OY 301 EAMLTSDLVMLPGMAOMGTVLGFLORGAOLHEFPQYKTHALATDPAFCALGEROG 360
 Db 301 EAMLTSDLVMLPGMAOMGTVLGFLORGAOLHEFPQYKTHALATDPAFCALGEROG 360
 OY 361 LEEDVVGORCPQDCITLQNTVSAGLNHQTFSVYAAYSVAQALNHTLQCNASGCPADP 420
 Db 361 LEEDVVGORCPQDCITLQNTVSAGLNHQTFSVYAAYSVAQALNHTLQCNASGCPADP 420
 OY 421 VKPQOLLENNMTNLFHVGLPLRPDSSGNVMEYDLKLMVQGSVPRLDVGRFNGSLRT 480
 Db 421 VKPQOLLENNMTNLFHVGLPLRPDSSGNVMEYDLKLMVQGSVPRLDVGRFNGSLRT 480
 OY 481 ERKTRMHTSN-----QKPVSRCSNOCQEGOVRRVKGHSCCYCVCDEASY 529
 Db 481 ERKTRMHTSN-----QKPVSRCSNOCQEGOVRRVKGHSCCYCVCDEASY 540
 OY 530 RQNPDDICTCGODEMSPERSTRCFRRSRFLANGEPVAVLLLLLSLALGVLAAGL 589
 Db 530 RQNPDDICTCGODEMSPERSTRCFRRSRFLANGEPVAVLLLLLSLALGVLAAGL 600
 OY 541 RQNPDDICTCGODEMSPERSTRCFRRSRFLANGEPVAVLLLLLSLALGVLAAGL 600
 Db 541 RQNPDDICTCGODEMSPERSTRCFRRSRFLANGEPVAVLLLLLSLALGVLAAGL 660
 OY 550 FVHRDSEPLVOASGGPLACFGLVCLSVLLEPPGSPARCIADQPLSHLPLTGCLIS 649
 Db 550 FVHRDSEPLVOASGGPLACFGLVCLSVLLEPPGSPARCIADQPLSHLPLTGCLIS 660
 OY 601 FVHRDSEPLVOASGGPLACFGLVCLSVLLEPPGSPARCIADQPLSHLPLTGCLIS 660
 Db 601 FVHRDSEPLVOASGGPLACFGLVCLSVLLEPPGSPARCIADQPLSHLPLTGCLIS 720
 OY 661 TLFLQAAEIFESELPSLMADRLSGCLRGPMAMLVLLAMVEVALCTWYLAPEPEVYT 720
 Db 661 TLFLQAAEIFESELPSLMADRLSGCLRGPMAMLVLLAMVEVALCTWYLAPEPEVYT 769
 OY 710 DWHMLPTALVHCRRSVWSEGLAHATNATLAFCLFTFLVRSQPGCYNRRARGLTFAML 769
 Db 710 DWHMLPTALVHCRRSVWSEGLAHATNATLAFCLFTFLVRSQPGCYNRRARGLTFAML 780
 OY 721 DWHMLPTALVHCRRSVWSEGLAHATNATLAFCLFTFLVRSQPGCYNRRARGLTFAML 780
 Db 721 DWHMLPTALVHCRRSVWSEGLAHATNATLAFCLFTFLVRSQPGCYNRRARGLTFAML 829
 OY 770 AFTITWSEFVPLLANVQVYLRPAVOMGALLCVLGIITLAAFLPRCYLLMRQPGMLTPEEF 829
 Db 770 AFTITWSEFVPLLANVQVYLRPAVOMGALLCVLGIITLAAFLPRCYLLMRQPGMLTPEEF 840

QY 830 IGGGPGDAGGNDGNTGNGKHE 852
 |||||||||||||||||||
 Db 841 IGGGPGDAGGNDGNTGNGKHE 863

RESULT 7
 AA080493
 ID AA080493 standard; Protein; 855 AA.

AC AA080493:

DT 12-MAR-2002 (first entry)

DE Human G-coupled receptor (GCREC) protein, Seq ID No 1.

XX Human: cytosolic; neuroprotective; immunosuppressant; nocitropic;
 KM anti-inflammatory; anti-viral; gastrointestinal; cardiovascular;
 KM cerebroprotective; G-coupled receptor; cell proliferative disease;
 KM lymphoma; leukemia; breast cancer; cirrhosis; neurological disorder;
 KM stroke; Alzheimer's disease; multiple sclerosis; mental retardation;
 KM cardiovascular disease; atherosclerosis; angina pectoris; indigestion;
 KM congestive heart failure; gastrointestinal disorder; dysphagia; AIDS;
 KM gastritis; autoimmune disorder; inflammatory disorder; Crohn's disease;
 KM systemic lupus erythematosus; metabolic disorder; diabetes; obesity;
 KM viral infection; herpesvirus; parvovirus;
 KM acquired immune deficiency syndrome.

OS Homo sapiens.

PN WO200190359-A2.

PD 29-NOV-2001.

XX 22-MAY-2001; 2001MO-US16833.

XX 22-MAY-2000; 2000US-206222P.

PR 23-MAY-2000; 2000US-207476P.

PR 02-JUN-2000; 2000US-208834P.

PR 07-JUN-2000; 2000US-208861P.

XX (INCY-) INCYTE GENOMICS INC.

XX Patterson C, Tribouley CM, Yao MG, Griffin JA, Thornton M, Lu Y;

PI Kallick DA, Gandhi AR, Au-Young J;

XX WPI: 2002-106199/14.

DR N-PSDB; ABR16615.

XX New G-protein coupled receptors useful for treating or preventing cell

PT proliferative (e.g. leukemia), neurological (e.g. stroke),

PT cardiovascular or autoimmune/inflammatory disorders -

XX Claim 1; Page 116-118; 148pp; English.

XX The invention relates to a novel human G-coupled receptor (I). (I) and
 CC its corresponding polynucleotides are useful for diagnosing, treating or
 CC preventing cell proliferative diseases (e.g. lymphoma, leukemia, breast
 CC cancer or cirrhosis), neurological disorders (e.g. stroke, Alzheimer's
 CC disease, multiple sclerosis or mental retardation), cardiovascular
 CC diseases (e.g. atherosclerosis, angina pectoris or congestive heart
 CC failure), gastrointestinal disorders (e.g. dysphagia, indigestion or
 CC gastritis), autoimmune/inflammatory disorders (e.g. AIDS, Crohn's disease
 CC or systemic lupus erythematosus) or metabolic disorders (e.g. diabetes
 CC or obesity), or viral infections (e.g. infection by herpesvirus or
 CC parvovirus). AA080493-AA080515 represent novel human G-coupled
 CC receptor amino acid sequences of the invention.

XX Sequence 855 AA;

Query Match 82.0%; Score 3708; DB 23; Length 855;
 Best Local Similarity 81.4%; Pred. No. 0;

Matches 738; Conservative 5; Mismatches 38; Indels 126; Gaps 10;

QY 1 MLGPAYGLSLMALHPGTGAPLCLSOQLRKKGQYVGLPEGEAREAGLRSTRSSP 60
 Db 1 MLGPAYGLSLMALHPGTGAPLCLSOQLRKKGQYVGLPEGEAREAGLRSTRSSP 60
 QY 61 VCTRFSSNGLLMALMKMAVEEINNKSDLLPGLTGYDLDPYCSSEPVYANKPSIMFLAKA 120
 Db 61 VCTRFSSNGLLMALMKMAVEEINNKSDLLPGLTGYDLDPYCSSEPVYANKPSIMFLAKA 120
 QY 121 GSRDIAVCNTYQYQPVYLAIVIGPHSSELMAYGKFFSFLMP----- 163
 Db 121 GSRDIAVCNTYQYQPVYLAIVIGPHSSELMAYGKFFSFLMP----- 163
 QY 164 -----OYSYGSMELLSARETFPSFRTVPDRQVLTAAELLOFGNNVAAL 212
 Db 164 -----OYSYGSMELLSARETFPSFRTVPDRQVLTAAELLOFGNNVAAL 212
 QY 213 GSDDVEYRGGLSTFSALAAARGICIAHEGLVPLPRADDSRLGRVQDVLHOVNOSSVOVLT 272
 Db 213 GSDDVEYRGGLSTFSALAAARGICIAHEGLVPLPRADDSRLGRVQDVLHOVNOSSVOVLT 272
 QY 241 GSDDVEYRGGLSTFSALAAARGICIAHEGLVPLPRADDSRLGRVQDVLHOVNOSSVOVLT 291
 Db 241 GSDDVEYRGGLSTFSALAAARGICIAHEGLVPLPRADDSRLGRVQDVLHOVNOSSVOVLT 291
 QY 273 LFASVHAHAFNYSISRSPKVVASSEAMVTSIDVYMGIPGMONGTVYGLFQRGQOLH 332
 Db 273 LFASVHAHAFNYSISRSPKVVASSEAMVTSIDVYMGIPGMONGTVYGLFQRGQOLH 332
 QY 292 LFASVHAHAFNYSISRSPKVVASSEAMVTSIDVYMGIPGMONGTVYGLFQRGQOLH 351
 Db 292 LFASVHAHAFNYSISRSPKVVASSEAMVTSIDVYMGIPGMONGTVYGLFQRGQOLH 351
 QY 333 EFPQYVTHLALATDPAPFCALGEREGLEEDVVGRCPOCDCTITLONVSAGLNHQTFSS 392
 Db 333 EFPQYVTHLALATDPAPFCALGEREGLEEDVVGRCPOCDCTITLONVSAGLNHQTFSS 392
 QY 352 EFPQYVTHLALATDPAPFCALGEREGLEEDVVGRCPOCDCTITLONVSAGLNHQTFSS 399
 Db 352 EFPQYVTHLALATDPAPFCALGEREGLEEDVVGRCPOCDCTITLONVSAGLNHQTFSS 399
 QY 393 VYAAVYSVAOLNHTLQCNASGCPADPVPKPMOLLENNYLTFFVHGLPLRFDSGGVNDM 452
 Db 393 VYAAVYSVAOLNHTLQCNASGCPADPVPKPMOLLENNYLTFFVHGLPLRFDSGGVNDM 452
 QY 400 -----RAQALHTLQCNASGCPADPVPKPMOLLENNYLTFFVHGLPLRFDSGGVNDM 452
 Db 400 -----RAQALHTLQCNASGCPADPVPKPMOLLENNYLTFFVHGLPLRFDSGGVNDM 452
 QY 453 EYDLKLMWQGSYPRLDVDFRNGSLTERLTKIMHNSD-----OKVSSCSR 501
 Db 453 EYDLKLMWQGSYPRLDVDFRNGSLTERLTKIMHNSD-----OKVSSCSR 501
 QY 453 EYDLKLMWQGSYPRLDVDFRNGSLTERLTKIMHNSD-----OKVSSCSR 512
 Db 453 EYDLKLMWQGSYPRLDVDFRNGSLTERLTKIMHNSD-----OKVSSCSR 512
 QY 502 QCGGQVRRVKGFSHCYDCVDCBAGSYRONPDIACTFCGODEMSPERSTRCFRRSRF 561
 Db 502 QCGGQVRRVKGFSHCYDCVDCBAGSYRONPDIACTFCGODEMSPERSTRCFRRSRF 561
 QY 513 QCGGQVRRVKGFSHCYDCVDCBAGSYRONPDIACTFCGODEMSPERSTRCFRRSRF 572
 Db 513 QCGGQVRRVKGFSHCYDCVDCBAGSYRONPDIACTFCGODEMSPERSTRCFRRSRF 572
 QY 562 LAMGEPAVLLLLLSLALGLVLAALGFVHHRSPLVQASGGGLACGLVCLGLVCLSV 621
 Db 562 LAMGEPAVLLLLLSLALGLVLAALGFVHHRSPLVQASGGGLACGLVCLGLVCLSV 621
 QY 573 LAMGEPAVLLLLLSLALGLVLAALGFVHHRSPLVQASGGGLACGLVCLGLVCLSV 632
 Db 573 LAMGEPAVLLLLLSLALGLVLAALGFVHHRSPLVQASGGGLACGLVCLGLVCLSV 632
 QY 622 ILFPQSPAPACLAQOPLSHPLTGCLSTFLQAAEFVSESLPLSWADRLSGCLRGPMA 681
 Db 622 ILFPQSPAPACLAQOPLSHPLTGCLSTFLQAAEFVSESLPLSWADRLSGCLRGPMA 681
 QY 633 ILFPQSPAPACLAQOPLSHPLTGCLSTFLQAAEFVSESLPLSWADRLSGCLRGPMA 692
 Db 633 ILFPQSPAPACLAQOPLSHPLTGCLSTFLQAAEFVSESLPLSWADRLSGCLRGPMA 692
 QY 682 WLVLVLAALVEVALCTVYLAAPPEVVTDMHMLPFEALVHCRTFSVWSFGLAAT----- 736
 Db 682 WLVLVLAALVEVALCTVYLAAPPEVVTDMHMLPFEALVHCRTFSVWSFGLAAT----- 736
 QY 693 WLVLVLAALVEVALCTVYLAAPPEVVT-----GLAHAAGGAG 731
 Db 693 WLVLVLAALVEVALCTVYLAAPPEVVT-----GLAHAAGGAG 731
 QY 737 -----NATLAFLCTGTFVLRSPG-----CYNRARGLFFAMAY 771
 Db 737 -----NATLAFLCTGTFVLRSPG-----CYNRARGLFFAMAY 771
 QY 732 ALPHLLTLLGQLRPSARHATATLCTETGF-----PGAEPAGPLQPCVASHICHAGLLHH 786
 Db 732 ALPHLLTLLGQLRPSARHATATLCTETGF-----PGAEPAGPLQPCVASHICHAGLLHH 786
 QY 772 FITWSEFVPLIAN-VQVVLBPAAVQMGALLLCVIGLIIAFLHPLRCYLIMRPGLTPEFFL 830
 Db 772 FITWSEFVPLIAN-VQVVLBPAAVQMGALLLCVIGLIIAFLHPLRCYLIMRPGLTPEFFL 830
 QY 787 --TGSHTVPLLAQAGSHSGPAVQMGALLLCVIGLIIAFLHPLRCYLIMRPGLTPEFFL 844
 Db 787 --TGSHTVPLLAQAGSHSGPAVQMGALLLCVIGLIIAFLHPLRCYLIMRPGLTPEFFL 844
 QY 831 GGGPGDA 837
 Db 831 GGGPGDA 851

RESULT 8

AAE10371

ID AAE10371 standard; Protein; 858 AA.

XX AAE10371.

XX AAE10371.

DT 10-DEC-2001 (first entry)
 XX Rat taste receptor, t1R3 protein.
 XX Rat, taste-cell-specific G protein-coupled receptor; t1R3; drug;
 KW genetic modulation; pharmaceutical; taste sensation; food industry;
 KW Chemosensory transduction.
 XX
 OS Rattus sp.
 XX
 PN W020016563-A2.
 PD 13-SEP-2001.
 PF 07-MAR-2001; 2001WO-US07265.
 XX
 PR 07-MAR-2000; 2000US-0187546.
 PR 07-APR-2000; 2000US-0195536.
 PR 06-JUN-2000; 2000US-0209840.
 PR 23-JUN-2000; 2000US-0214213.
 PR 17-AUG-2000; 2000US-0226448.
 PR 03-JAN-2001; 2001US-0259227.
 XX
 PA (SENO-) SENOMTX INC.
 PI Adler JE, Zozulya S, O'Connell SM, Li X, Staszewski LJ;
 DR N-PSDB; AAD17514.
 XX
 PT New mammalian taste-cell-specific G protein-coupled receptor
 PT polypeptides for identifying compounds that modulate taste signaling
 PT are useful in food, to modulate the sweet taste of foods or drugs -
 PS Claim 152; Page 77; 119pp; English.
 XX
 CC The invention relates to mammalian taste-cell-specific G protein-coupled
 CC receptors, t1R and their corresponding cDNA molecules. Taste receptors,
 CC t1R are useful for screening compounds which are used to activate or
 CC modulate chemosensory transduction, such as taste sensation. The
 CC identification and isolation of novel taste receptors and taste
 CC signalling molecules allow for new methods of chemical and genetic
 CC modulation of taste transduction pathways. The taste modulating
 CC compounds are useful in pharmaceuticals and food industries to improve
 CC the taste of a variety of consumer products, or to block undesirable
 CC tastes, e.g., in certain pharmaceuticals. t1R's are also useful in
 CC biochemical assay for identifying tastant (t1R) ligands having binding
 CC specificity for t1R involved in taste signalling. The present sequence is
 CC rat taste-cell-specific G protein-coupled receptor, t1R3 protein.
 XX
 SQ Sequence 858 AA:
 Query Match 71.9%; Score 3255; DB 22; Length 858;
 Best Local Similarity 72.9%; Pred. No. 2.6e-304;
 Matches 623; Conservative 74; Mismatches 144; Indels 14; Gaps 4;
 QY 1 MGPVAVLGSLMALHPGAPLCLSQOLRMKGDVYLGFLPGAEEAGLRPRPSP 60
 DB 1 MGLAVIIGSLAFLFELGSSSLCSQOKRAGDVIILGFLFLGTTEETLQNRQPKNI 60
 QY 61 VCTRSSNGLLMALAKMAVEEINKSDLLPGLRLGYDLFDTCSEPVAMKPSLFLAKA 120
 DB 61 LCTRESPLGLFLAMAKMAVEEINNGSALLPGLRLGYDLFDTCSEPVYTMKPSLFLMAKV 120
 QY 121 GSRDIAAYCNTQYOPRVLAIVGPHSSSELAMTYGKFFSFLMPQVSYGSMELSLARETF 180
 DB 121 GSQSIATAYCNTQYOPRVLAIVGPHSSSELALITGKFFSFLMPQVSYGSMELSLARETF 180
 QY 181 PSFFFTVPSPDRYQVLAELLOEFGNNVVAALGSDDEYRGOSLIFSALAAAGCICIAHE 240
 DB 181 PSFFFTVPSPDRYQVLAELLOEFGNNVVAALGSDDEYRGOSLIFSALAAAGCICIAHE 240
 QY 241 GLVPLPRADSRLLGKVVQVLYHVNOSVQVLLFASVHAHALFNYSISSRLSPKVVAS 300

DB 241 GLVPLPRADSRLLGKVVQVLYHVNOSKVVYLLFASARAVYSLSFVSLIHDLSPKVWAS 300
 QY 301 EAMLSIDLVMGLPGMAQMTVIGFLORGAQLHEFPQYKTHIALATDPAGSALGEREG 360
 DB 301 ESMLSIDLVMGLPGMAQMTVIGFLORGAQLHEFPQYKTHIALATDPAGSALGEREG 360
 QY 361 LEEDVVGORCPQODCTTQNVASGLN-----HHQFSVYAAYSVQAALHNTLQCA 412
 DB 360 LEEDVVGORCPQODCTTQNVASGLN-----HHQFSVYAAYSVQAALHNTLQCA 412
 QY 413 SCGPADDPYKPMQLENNMNTLFHVGLPLRFDDSGNVDMEDYDKLMVQSGVPRLDHVG 472
 DB 420 SHCHTSEPVQPMQLENNMNTLFHVGLPLRFDDSGNVDMEDYDKLMVQSGVPRLDHVG 479
 QY 473 RFRNGSLRTERLKRIMHTSDNQKPVSRCSQCOBGVRYKGFHSCCYDCVCEAGSYRON 532
 DB 480 RFRNGSLRTERLKRIMHTSDNQKPVSRCSQCOBGVRYKGFHSCCYDCVCEAGSYRON 537
 QY 533 PDICTCTPCGQDMSPERSTRCFRRSRPLANGEPVALLLLSLALGLVLAAGLFPVH 592
 DB 538 PDICTCTPCGQDMSPERSTRCFRRSRPLANGEPVALLLLSLALGLVLAAGLFPVH 597
 QY 593 HRDSPLOVQASGGPLACGLVCLGLVCLSVLLEPGQSPARCIAQDPLSLPTGCLSTLF 652
 DB 598 YWDSPLVQASGGPLACGLVCLGLVCLSVLLEPGQSPARCIAQDPLSLPTGCLSTLF 657
 QY 653 LQAAELFVSELPPLSWADLSGCLKRPAMLYVLLAMLEVACTWYLVAFPEVYTDH 712
 DB 658 LQAAELFVSELPPLSWADLSGCLKRPAMLYVLLAMLEVACTWYLVAFPEVYTDH 717
 QY 713 MPTREALVCRFRSWSFGLAHATNATLAFCLFGLFELVRSQGCYNRARGLFMALAYF 772
 DB 718 VLPTEVLEHCRMRWSVSLGVIHTNVAFLFCLFGLFELVRSQGCYNRARGLFMALAYF 777
 QY 773 TIWVSFVPLLANVQVYVLRPAVQMGALLLCVIGILAAFHLPRLCYLRLMRDGLMTPFEFLGG 832
 DB 778 TIWVSFVPLLANVQVYVLRPAVQMGALLLCVIGILAAFHLPRLCYLRLMRDGLMTPFEFLGG 837
 QY 833 GEGDAQGQDGNITGN 847
 DB 838 SPKEA---SDGNSS 849
 RESULT 9
 AAU73183
 ID AAU73183 standard; Protein: 858 AA.
 AC AAU73183;
 XX
 DT 12-MAR-2002 (first entry)
 XX
 DE Murine SACL polypeptide.
 KW Human; mouse; SACL; carbohydrate; sweetener; ethanol; alcoholism;
 KW obesity; diabetes; transgenic embryo; body tissue; body fluid; pancreas;
 KW blood; tongue; anorectic; antidiabetic; gene therapy;
 KW protein replacement therapy.
 OS Mus sp.
 XX
 PN W0200183749-A2.
 PD 08-NOV-2001.
 PF 25-APR-2001; 2001WO-US13387.
 XX
 PR 28-APR-2000; 2000US-2007949.
 PR 28-JUL-2000; 2000US-221419P.
 PR 10-NOV-2000; 2000US-247443P.
 XX
 PA (WARN) WARNER LAMBERT CO.
 PA (MONE-) MONELL CHEM SENSES CENT.

CC Infarction, ulcer, inflammation or benign prostatic hypertrophy
 CC and infections (e.g. bacterial, fungal, protozoan or viral).
 CC Other diseases that may be treated include Alzheimer's disease,
 CC Parkinson's disease, dementia, diabetes, leukemia and lymphoma.
 CC Sequences of the invention are used for identifying agonists and
 CC antagonists that modulate their activity. They are also used in
 CC gene therapy. The present sequence is human GPCR protein.

XX Sequence 596 AA;

Query Match 68.9%; Score 3116; DB 23; Length 596;
 Best Local Similarity 89.8%; Pred. No. 3.8e-291;
 Matches 596; Conservative 0; Mismatches 0; Indels 68; Gaps 1;

QY 164 QVSYGASMEILLASRETFPSFRTPVSDRVOLTAALAEILOEFGMVVAALGSDDEFGROGL 223
 DB 1 QVSYGASMEILLASRETFPSFRTPVSDRVOLTAALAEILOEFGMVVAALGSDDEFGROGL 60
 QY 224 SIFSALAAARGICIAHEGLVPLPRADSRIGKVDVLIHQVQSSVQVVLLEFASVHAHAAL 283
 DB 61 SIFSALAAARGICIAHEGLVPLPRADSRIGKVDVLIHQVQSSVQVVLLEFASVHAHAAL 120
 QY 284 FNYGISISRLSPKVVAVASAMLTSDLYMGLPGMAQMGTVLGLQKGAOLHEFPQYKTHLA 343
 DB 121 FNYGISISRLSPKVVAVASAMLTSDLYMGLPGMAQMGTVLGLQKGAOLHEFPQYKTHLA 180
 QY 344 LATDPAFCSALGEREOGLEEDVVGORPCDITLQNVSAGLNHQFESVAAAYSAVQA 403
 DB 181 LATDPAFCSALGEREOGLEEDVVGORPCDITLQNVSAGLNHQFESVAAAYSAVQA 240
 QY 404 LHNTLQCNASGCPADDPVKPMOLLENNYNIHFVHGLPLRPDSQSNVDMEXDLKIMVWG 463
 DB 241 LHNTLQCNASGCPADDPVKPM----- 261
 QY 464 SVPLHVDYGRNGSLRTERLKIKNHTSDNQPVSRCSQCQEGVRRYKGFHSCCYD 523
 DB 262 -----QKPVSRCSQCQEGVRRYKGFHSCCYD 292
 QY 524 CEASGYRNPDDIACFCGODEMSPERSTRCFRRRSRFLANGEPVLLLLLSTALGLV 583
 DB 293 CEASGYRNPDDIACFCGODEMSPERSTRCFRRRSRFLANGEPVLLLLLSTALGLV 352
 QY 584 LAAIGLVNHRDSPLYQASGGPLACFGLVCLGLVLSVLEPPGPPSPARCLAQOPLSHLP 643
 DB 353 LAAIGLVNHRDSPLYQASGGPLACFGLVCLGLVLSVLEPPGPPSPARCLAQOPLSHLP 412
 QY 644 LTGLSTLFLQAAELFVSESLPLSMADRLSGCLNGPMAVLYVLLAMLYEVALCTWYLVAF 703
 DB 413 LTGLSTLFLQAAELFVSESLPLSMADRLSGCLNGPMAVLYVLLAMLYEVALCTWYLVAF 472
 QY 704 PPEVVTDMHMLPTEALVHCRTSMWSEFGLAHATNATLAFCLGTFELVRSQPGCYNRAG 763
 DB 473 PPEVVTDMHMLPTEALVHCRTSMWSEFGLAHATNATLAFCLGTFELVRSQPGCYNRAG 532
 QY 764 LTFPMLAYFITWVSFVPLLANVQVLLRPAYOMGAILLCVIGITLAAPHILPCYLLMRQGL 823
 DB 533 LTFPMLAYFITWVSFVPLLANVQVLLRPAYOMGAILLCVIGITLAAPHILPCYLLMRQGL 592
 QY 824 NTPE 827
 DB 593 NTPE 596

RESULT 11

AAE21817 ID AAE21817 standard; Protein: 370 AA.

AAE21817;

16-JUL-2002 (first entry)

Human GPCR protein #1.

KW Human; G protein-coupled receptor; GPCR; central nervous system disease;
 KW cancer; anorexia; bulimia; asthma; benign prostatic hypertrophy; ulcer;
 KW allergy; acute heart failure; myocardial infarction; urinary retention;
 KW Parkinson's disease; Alzheimer's disease; hypotension; angina pectoris;
 KW hypertension; osteoporosis; inflammation; infection; dementia; diabetes;
 KW leukaemia; lymphoma; gene therapy.

OS Homo sapiens.

FX Key Location/Qualifiers
 FT Region 35..61
 FT /note- "G-protein_recep_F3-2 region"

PN W0200224885-A2.

PD 28-MAR-2002.

PF 21-SEP-2001; 2001WO-EP10929.

PR 25-SEP-2000; 2000US-234982P.

PA (FARB) BAYER AG.

P1 Kossida S;

DR WPI: 2002-339949/37.

XX N-PSDB: AAD34458.

PT New G-protein coupled receptor polynucleotide and its encoded protein,
 PT useful for identifying modulators of GPCR activity, and in gene therapy
 PT for treating bacterial infection, cancer, acute heart failure or
 PT inflammation

PS Claim 25; Fig 2; 101pp; English.

XX The invention relates to human G protein-coupled receptor (GPCR)
 CC polypeptides and polynucleotides. GPCR sequences are useful for
 CC preventing or ameliorating the diseases such as cancer, anorexia,
 CC bulimia, asthma or other allergies, peripheral or central nervous
 CC system disease, acute heart failure, hypotension, hypertension,
 CC urinary retention, osteoporosis, angina pectoris, myocardial
 CC infarction, ulcer, inflammation or benign prostatic hypertrophy
 CC and infections (e.g. bacterial, fungal, protozoan or viral).
 CC Other diseases that may be treated include Alzheimer's disease,
 CC Parkinson's disease, dementia, diabetes, leukemia and lymphoma.
 CC Sequences of the invention are used for identifying agonists and
 CC antagonists that modulate their activity. They are also used in
 CC gene therapy. The present sequence is human GPCR protein.

XX Sequence 370 AA;

Query Match 43.0%; Score 1944; DB 23; Length 370;
 Best Local Similarity 100.0%; Pred. No. 2.1e-178;
 Matches 360; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 493 QKPVSRCSQCQEGVRRYKGFHSCCYDVCDEASGYRNPDDIACFCGODEMSPERST 552
 DB 11 QKPVSRCSQCQEGVRRYKGFHSCCYDVCDEASGYRNPDDIACFCGODEMSPERST 70

QY 553 RCFRRRSRFLANGEPVLLLLLSTALGLVLAALGLFVNHRDSPLYQASGGPLACFGLV 612
 DB 71 RCFRRRSRFLANGEPVLLLLLSTALGLVLAALGLFVNHRDSPLYQASGGPLACFGLV 130

QY 613 CLGIVCLSVLPPGPPSPARCLAQOPLSHLPITGCLSTLFLQAAELFVSESLPLSMADRL 672
 DB 131 CLGIVCLSVLPPGPPSPARCLAQOPLSHLPITGCLSTLFLQAAELFVSESLPLSMADRL 190

QY 673 SGCLNGPMAVLYVLLAMLYEVALCTWYLVAFPEVVTDMHMLPTEALVHCRTSMWSEFGL 732
 DB 191 SGCLNGPMAVLYVLLAMLYEVALCTWYLVAFPEVVTDMHMLPTEALVHCRTSMWSEFGL 250

QY 733 AHATNATLAFCLGTFELVRSQPGCYNRAGLTFPMLAYFITWVSFVPLLANVQVLLRPA 792
 DB 733 AHATNATLAFCLGTFELVRSQPGCYNRAGLTFPMLAYFITWVSFVPLLANVQVLLRPA 792

OY 818 MROPELNTPEFF 829
 Db 1127 LCRPELNTPEHF 1138

RESULT 13
 ID AAY45028 standard; Protein: 840 AA.

AC AAY45028;
 DT 31-MAY-2000 (first entry)

DE Rat sensory transduction G-protein coupled receptor-B3 variant #3.
 XX
 XX Rat; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;
 KW Sensory cell; taste receptor cell; screen; taste modulator;
 KW pharmaceutical; food industry; taste topographic map; tongue; variant.
 XX
 OS Rattus sp.
 XX Synthetic.

XX Key Location/Qualifiers
 FH Misc-difference 90
 FT /note="Wild type Ala substituted by Gly"

XX W0200006592-A1.
 XX 10-FEB-2000.
 XX 27-JUL-1999; 99WO-US17099.
 XX 28-JUL-1998; 98US-0094465.
 XX (REGC) UNIV CALIFORNIA.
 PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX
 PI Zuker CS, Adler JE, Lindemeyer J, Ryba N, Hoon M;
 DR WPT; 2000-205451/18.

XX New isolated sensory transduction G-protein coupled receptor, useful
 PT for developing products for use in studying and modulating the taste
 PT transduction pathway -
 PS Disclosure; Page -: 83pp; English.

XX The present sequence is a polymorphic variant #3 of rat taste cell
 CC specific G-protein coupled receptor, GPCR-B3 which is involved in
 CC sensory transduction. GPCR-B3 is specifically expressed in foliate
 CC and fungiform cells, with lower expression in circumvallate taste
 CC receptor cells of the tongue. The protein has an extracellular domain,
 CC seven transmembrane domains and an intracellular domain.
 CC The GPCR-B3 sequence is used to screen compounds that modulate sensory
 CC signalling in taste cells, especially taste modulators useful in
 CC pharmaceutical and food industries to customise taste. The sequence
 CC can also be used as probe for identifying taste cells and
 CC subsets of taste receptor cells such as foliate, fungiform and
 CC circumvallate. Such probes are also useful to generate taste
 CC topographic maps that elucidate the relationship between the taste
 CC cells of the tongue and sensory neurons leading to taste centres
 CC in the brain.
 CC Note: The present sequence is not given in the specification but is
 CC derived from rat GPCR-B3 sequence shown in page 75 (AAY45021).
 XX
 XX Sequence 840 AA;

Query Match: 26.7%; Score 1209.5; DB 21; Length 840;
 Best Local Similarity 33.8%; Pred. No. 5,5e-107;
 Matches 291; Conservative 151; Mismatches 339; Indels 79; Gaps 18;
 OY 12 WALLHPTGAPLCLSQOL-----RMKGDIYVLGGGLPPL-GEAEERAGLR 53

Db 4 WA-----AHLLSLQLYVCMAPSCQRTSESSPGRSLPDDFLAGLFLSHDDC-----LQV 52
 OY 54 RTRDSSPVCTR--FSSNGLMALAMKMAVEINKSDLLPGLRGYDLPJTCSPEVAM 110
 Db 53 RHRPLVTSQDRPDSENGHYHFLQMRFTVEIINNSSLGILPNITGLYELDYVCSSES-ANV 111
 OY 111 KPSLMFLAKAGSRDIAACNTQYQPRVLAIVIGPSSSELAWTGKFFSFLMPQVSGAS 170
 Db 112 YATLKVIALQGRHTEICKDRNHSKVAFIPDNDTHAATTAALCPGLMPLVSYRAS 171
 OY 171 MELLSARETFPFFEFRTYVSDRVOLTAABELLOEFSGMNVAAALSDDEXROGLSIFSA 230
 Db 172 SVLTSARKRFPSEFLRTVSDRHQVEVMYQLQSFEMVWVMSLIGSYGDIQGLVQALBELA 231
 OY 231 AARGICIAHEGLVPLPRADSRIG--KVQDYLHOVNOSSVOVILLFASVHAHALENTSI 288
 Db 232 VPRGICVAFKDIVP-----SARVGDPRMQSMQHLQARTVYVYVFSNRHLARFFRSV 287
 OY 289 SSRSPKVMVASEAMLTSDLVMGLPGMQMGTVLGFQRGAO---LHFEPO-YVKTAL 344
 Db 288 LANLTGKVMVASEDMALSTYITSTVIGIGIVLGVAVQORVGLKFEESYVAATAA 347
 OY 345 AT---DPAFCALGEREGLDEDVVGRCPODCITLQNVS--AGLNHQTFSVYAAVYS 399
 Db 348 PSACEGSMCS-----TNOLCRCHFTTTNNMPLTIGAFMSAAVRYVEAYYA 394
 OY 400 VAQALHNTLQCNASGCPADPVKRWQLLENMNTLFHVGLPLRPDSGAVDMEYDLK 459
 Db 395 VAHGHLQGLGCTSEIC-SRGVYYPWQLIQYKVFLLHENTVAVDDNDGTLYGYYDIAM 453
 OY 460 VMQGSVPRLDHVGRENDS---LTERLKIHMHTSDNOKPVSRCSRCQCGQVRRYKGP 516
 Db 454 DMNGPEMTFETIGASLSPVHLIDINKTKTIOMHKNQVPASVCTDCLAGHRRVYVSGH 513
 OY 517 CCYCDVDEAGSYRONPDDIACTEGODEWSPERSTRCFRRRSRLRANGEPAYVILL 576
 Db 514 CCFECVPCFAGFLMMSLHLIQPCGTEDMAPKRESTCFPRVRELAMHEPISLVLIAN 573
 OY 577 SLALGLVLAIGLPHVHNDSPVQASGCPACFGVLCGLVCLSTVLLPQGPSAPARCLAQ 636
 Db 574 TLLLLLVGTAGLPMHHTPVRSAGRLCFMLGSLVAPSSCSYSPFGEPYAPACLLR 633
 OY 637 QPLSHLPFGCLSTFLQAAELFV---ESELPL---SMADRLSCLGPMAMLVLLAM 689
 Db 634 QPLSLGFAIRPLSCLTISFQLVIFKSTKVPTRYRMA-----QNRHAGLFVIYSS 686
 OY 690 LVEVALCTWYLVAFPPPEVYTDHMLPTBALVHCRTSRWVSFGLAHATNATLAFICLGT 749
 Db 687 TVHLILCTLWLVMTPRPTREYQRPHLVILECTEVNSVGFLLAFTNILLISTFVCSY 746
 OY 750 LVRSQPGCYNARAGLTFMLAVFTWVSFVPLNAVQVLLPRAYOMGALLCVIGLILAF 809
 Db 747 LKELPENYINAKCYTFELNLNPFVMAIFFTWASLYQSYLPAVAVLAGLTLTLGSGFG 806
 OY 810 HLPRCYILMRPGJLNTPEFF 829
 Db 807 FLPRCYVILCRPELNTPEHF 826

RESULT 14
 ID AAY45021 standard; Protein: 840 AA.

AC AAY45021;
 DT 31-MAY-2000 (first entry)

DE Rat sensory transduction G-protein coupled receptor-B3.
 XX
 XX Rat; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;
 KW Sensory cell; taste receptor cell; screen; taste modulator;
 KW pharmaceutical; food industry; taste topographic map; tongue.

```

XX OS Rattus sp.
XX Key Location/Qualifiers
XX FH Key 1..580
XX FT Domain /label= Extracellular_domain
XX PN WO200006592-A1.
XX PD 10-FEB-2000.
XX
XX PF 27-JUL-1999; 99WO-US17099.
XX PR 28-JUL-1998; 98US-0094465.
XX
XX PA (REGC ) UNIV CALIFORNIA.
XX PI (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX PI Zuker CS, Adler JE, Lindemeier J, Ryba N, Hoon M;
XX DR WPI; 2000-205451/18.
XX DR N-PSDB; AA250743.
XX
XX PT New isolated sensory transduction G-protein coupled receptor, useful
XX PT for developing products for use in studying and modulating the taste
XX PT transduction pathway.
XX PS
XX PS Claim 22; Page 75; 83pp; English.
XX
XX CC The present sequence is a taste cell specific G-protein
XX CC coupled receptor, GPCR-B3 which is involved in sensory transduction.
XX CC This sequence was isolated from the 12AP rat circumvallate cDNA library.
XX CC GPCR-B3 is specifically expressed in foliate and fungiform cells, with
XX CC lower expression in circumvallate taste receptor cells of the tongue.
XX CC The protein has an extracellular domain, seven transmembrane
XX CC domains and an intracellular domain.
XX CC The present sequence is used to screen compounds that modulate sensory
XX CC signalling in taste cells, especially taste modulators useful in
XX CC pharmaceutical and food industries to customize taste. The sequence
XX CC can also be used as probe for identifying taste cells and
XX CC subsets of taste receptor cells such as foliate, fungiform and
XX CC circumvallate. Such probes are also useful to generate taste
XX CC topographic maps that elucidate the relationship between the taste
XX CC cells of the tongue and sensory neurons leading to taste centres
XX CC in the brain.
XX CC
XX SQ Sequence 840 AA;

```

```

Query Match 26.7%; Score 1208.5; DB 21; Length 840;
Best Local Similarity 33.8%; Pred. No. 6.9e-107;
Matches 291; Conservative 151; Mismatches 339; Indels 79; Gaps 18;

```

```

OY 12 WALHPGTGAPICLSOOL-----RMKGDVIVGGLPL-GEAEAGIRLS 53
DB 4 WA-----AHLISLQLVCMASFSCORTSESSPEFSLPDGLAGLSLHSDC---LQV 52
OY 54 RTRESSEVCTR---FSSNGILMALAMKAVEEINKKSDLLPGLRIGYDLPDTCSEPVVAM 110
DB 53 RHRRLVTSCDRPSDFNGHGLFQAMFTVEINNSALLPNTLIGVELDYCES-ANY 111
OY 111 KPISMLFAKAGSRDIAYCITYQPRYLAVIGPHSSELMATVTKFSSFLMPQVSGAS 170
DB 112 YATIRVIALOGPRHIELOKDLRNHSSKVAFIGPDNTDHAVTFAALLGPFMLPLVSYEAS 171
OY 171 MELLSARETFSPFRVPSDVOLTAALLOERGMVMAVAGSDDEYGRQGLSIFPALA 230
DB 172 SVVLSARKRPFSLRIVPSDRHOVEYVQLOSGFWMISLIGSYGYGLGVQVLALELA 231
OY 231 AARGCIIAHGLVPLPRADSLRG--KVQDVLIQVNOSSVQVVLIFASVNAHAHLFYNSI 288
DB 232 VPRGICVAFDIYDF-----SARVGDPRMQSMQMLAQARTTVVVFENRHLARFFESV 287
OY 289 SSRISPFWVWVASEAMLTSDLVNGLPGMAQMGVIGFLQGAQ---LHEFPO-YVKTHIAL 344

```

```

DB 288 LANLTGKVVWVASEDMALSTYITSVTGIOGIGYVGAVVOORQVPGLEFESEYRAVTA 347
OY 345 AT---DPAFCSALGERDGLIEDVVGORCPQDCITLQNS--AGLNHHQFSSYAAVVS 399
DB 348 PSACESSWCS-----TNOLCRECHFTTRNNPDLGAFSMAAIFYEAVYA 394
OY 400 VQALHNTLQCNAGCPADPVKPMOLLENNYNLTFHVGGIPLRFDSGGVMDYDLKLW 459
DB 395 VAGHLHQLIGTSEIC-SRGPVYPMQLQOYKVNFLHENTVAFDNDGTLGYDILAW 453
OY 460 WQGSVRLHDVGRFNS---LTERLKRHNHSDNKPVSRCRQCEGVRRVYKGFHS 516
DB 454 DWNPEWTFELIGSASLSPVHLIDINKRKIOWHGKNNQVPSVCTDIDLAGHRRVVGSH 513
OY 517 CCYVCVDCGAGSYRQNDLACTFCGODEMSPERSTCFRRRSFLAMGEPVALLLLL 576
DB 514 CCEFCVPCGEGTFLNMSSELHICOPCGTEWAPKESSTCFPTVEFLAMHEPISLVLTAN 573
OY 577 SLALGLVLAALGLFVHHRDSPVQASGGLACFGLVCLGLVCLSVLLPQGPSPARCLAQ 636
DB 574 TLILLILVGTAGLFAWHFHPVVRASAGRLCTMLGSLVAGSCSFYFPGEPVPAQLLR 633
OY 637 QPLSHPLVTCGLSTFLQAAEFV----ESELPL---SMADRLSGCLRGPMALVLIAM 689
DB 634 QPLFSLGFAIFLSCLTIRSFOLVIFKFSRKVPFRRTWA-----QNHGAGLFIYSS 686
OY 690 LVEVALCTWYLVAPPEVVDHMLPTEALVHCTREMSVSGILAHATNATLAFCEFTF 749
DB 687 TWHLLICTWLVMMTPPTREYORFPLVILECEVSVGLFTNILLISTFVCSY 746
OY 750 LVRSOPCYNARNARGLAFMLAFPTWVSFVPLANVOVLRPAVOMKALLCVGLIAAF 809
DB 747 LGRLEPENVYNAKCVTRSSLINFEVSWIAFTMASIYGSTLPVAVNLAAGLTLISGFSGY 806
OY 810 HLPKCYLLAMPGLNTPPEF 829
DB 807 FLPCYVILCRPELNTEHF 826

```

RESULT 15

AAV45026
ID AAV45026 standard; Protein; 840 AA.

AC AAV45026;

DT 31-MAY-2000 (first entry)

DE Rat sensory transduction G-protein coupled receptor-B3 variant #1.

KW Rat; sensory transduction; G-protein coupled receptor-B3; GPCR-B3;
KW sensory cell; taste receptor cell; screen; taste modulator;
KW pharmaceutical; food industry; taste topographic map; tongue; variant.

OS Rattus sp.
XX Synthetic.

XX Key Location/Qualifiers

XX MISC-difference 33 /note="Wild type Leu substituted by Ile"

XX WO200006592-A1.

XX 10-FEB-2000.

XX 27-JUL-1999; 99WO-US17099.

XX 28-JUL-1998; 98US-0094465.

XX (REGC) UNIV CALIFORNIA.

XX (USSH) US DEPT HEALTH & HUMAN SERVICES.

XX Zuker CS, Adler JE, Lindemeier J, Ryba N, Hoon M;

```

XX  WPI: 2000-205451/18.
DR
XX  New isolated sensory transduction G-protein coupled receptor, useful
PT  for developing products for use in studying and modulating the taste
PT  transduction pathway
PS  Disclosure: Page -, 83pp; English.
XX
CC  The present sequence is a polymorphic variant #1 of rat taste cell
CC  specific G-protein coupled receptor, GPCR-B3 which is involved in
CC  sensory transduction. GPCR-B3 is specifically expressed in foliate
CC  and fungiform cells, with lower expression in circumvallate taste
CC  receptor cells of the tongue. The protein has an intracellular domain,
CC  seven transmembrane domains and an intracellular domain.
CC  The GPCR-B3 sequence is used to screen compounds that modulate sensory
CC  signalling in taste cells, especially taste modulators useful in
CC  pharmaceutical and food industries to customise taste. The sequence
CC  can also be used as probe for identifying taste cells and
CC  subsets of taste receptor cells such as foliate, fungiform and
CC  circumvallate. Such probes are also useful to generate taste
CC  topographic maps that elucidate the relationship between the taste
CC  cells of the tongue and sensory neurons leading to taste centres
CC  in the brain.
CC  Note: The present sequence is not given in the specification but is
CC  derived from rat GPCR-B3 sequence shown in page 75 (AAY45021).
XX
SQ  Sequence 840 AA:

Query Match          26.7%, Score 1208; DB 21; Length 840;
Best Local Similarity 33.9%; Pred. No. 7.7e-107;
Matches 291; Conservative 151; Mismatches 342; Indels 74; Gaps 19;

OY  6 VLGLSL---WAL-----LHPTGAPLCLSQGLRMKGDYVGLGFPPL-GEAEFAGLRRT 55
DB  9 LLSLDLVYCMASCSQRTSSPGEFSLP-----GDFLAGLFLSLHGDG---LQVNH 54
OY  56 RPSSEVCTR---FSSNGLLMALMAKMAVEEINKSDLLPGLRIAGYDLDFTCSEPVYAMKP 112
DB  55 RPLVYSCDRPDSFNGHGLHQMRFYVEEINSSALLPNTILGLEYIDVCS-ANYYA 113
OY  113 SLMFPAKAGSRDIAYCINTQYQPRVLAIVGPHSELMAVTGKFFSFLMPOVSYGASME 172
DB  114 TLRVLAOGPRHIEIQKLRHSSKVAFAIGPDNDHAYTTAALLGPFLMPLVSYEASSV 173
OY  173 LLSARETPESEFRVPSDVOYLTAAELTQERGMWVAALGSDDEYRGQLSIFSAALAA 232
DB  174 VLSARKKPPSLRTVPSDHOEVNVOQLDSFGWVWISLISYGDYGOALAEELAVP 233
OY  233 RQICIAHGLVPLPRADSRLLG--KVQDVLHVNOSVOVVLFFASVNAHALFNYSSIS 290
DB  234 RQICVAFDIYDF---SARVGDPRMQSMQMLAQARTTVVVFNSRHILARVFFRSVILA 289
OY  291 RLSPKVVWASEAWLTSDLVNGLPGMAQMGTYIGFLQGAQ---LHEPQ-YVKTHLALAT 346
DB  290 NLTGVVWVASEMAISTYITSTYTGIGCTVGVAVOORVPGLEKEFESYRAVTAAPS 349
OY  347 ---DPAFCSALGERQGLEEDVVGORPCOCITLONVS--AGLNHQTFSYAAVYSA 401
DB  350 ACPBSGMS------TNOLCRCHTFTTRNMPILGAFMSAAYRYEAIVAYA 396
OY  402 QALAHTLQCNASGCPADDPVYKQWLENNYNTLFHVGGLPLRFDSGQVNDMEYDLKLVW 461
DB  397 HGLHQLLCTSEIC-SRGPVYPMQLQIYKVNFLHENTVAFDNGDILGYDIIANDW 455
OY  462 QGSVRLHDVGRFNGS---LRTERLKIRNHTSDNOKPVSRCSQCEGOVRRVYKGFHSQC 518
DB  456 NGPEMTFELIGSASLSPVHLIDINKTKIQHGRKNQVPVSVCTTDLAGHHFVYVSSHCC 515
OY  519 YDCVCEAGSYRQNDPDLACFCGDEWSPERSTCFRRRSRFLAMGEPVAVLLLLLSTL 578
DB  516 FECVCEAGSTFLNMSELHIQPCGTEWAPKESTCTCFPTVEFLAMHEPISLVLIAANTL 575

```

```

OY  579 ALGLVLAAGLFVHHRDSPLVOASGGPLACFGVLCGLVCLSVLLFPGOPSPARCLAOOP 638
DB  576 LLLLVGTAGLFAMHFHTPVVRSAGRCFLMLGSLVAGSCSFYSFFGEPTVPACLLAQP 635
OY  639 LSHPLTGLSTLFIQAAEIFV---ESELPL---SWADRUSGCLRGFWAMLVYLLAMLY 691
DB  636 LFSLGFAFLFSLQTLIRSFQLVLIIFKFSYKVPFVRYTMA-----QNHGAGLFIYSSTV 688
OY  692 EVALCTWLVAFPPREVYVDMHMLPTREALVHCRTSRWSVFGLAHATNATLAPLCFLGTFLV 751
DB  689 HLLICLTVMWTPTPTREYQRPFLVLIELCTEVNSVGFLLAFTHNILLISTFVCSYLG 748
OY  752 RSQPCYRARGLFEAMLAYFITWVSFVPLLANOVVLRPAVOMGALLCYGLIAAPHL 811
DB  749 KELPENYNKACVPTSSLNLFVSWIAFTMASTYQGSYLPANVLAAGLITLSGFSGYFL 808
OY  812 PCYVILMROPGLNTPPEFF 829
DB  809 PKCVYILCRPELINTNEHF 826

```

Search completed: May 19, 2003, 09:50:04
Job time : 42.3065 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5_4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:47:40 ; Search time 20.1657 Seconds

(without alignments)
4061.679 Million cell updates/sec

Title: US-09-927-315-15

Perfect score: 4524

Sequence: 1 MGPVAVLGSLMALHPGCTG.....GPGDAQGQNDGNTGNGKHE 852

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283224 seqs, 96134422 residues

Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3229	71.4	858	2	JC7683
2	1068.5	23.6	1078	2	AC6715
3	1060	23.4	1085	2	SA40476
4	1057.5	23.4	1088	2	B56715
5	1054	23.3	1079	2	159362
6	670.5	14.8	872	2	JH0561
7	631.5	14.0	871	2	A46742
8	619.5	13.7	912	2	JH0563
9	612.5	13.5	1218	2	571376
10	604.5	13.4	915	2	A49874
11	597.5	13.2	879	2	JH0562
12	593.5	13.1	908	2	I49142
13	592.5	13.1	879	2	JC7160
14	584	12.9	1171	2	A42916
15	580.5	12.8	1180	2	JC2132
16	580.5	12.8	1212	2	JC2131
17	574.5	12.7	1199	2	A41939
18	474	10.5	999	2	T21628
19	436.5	9.6	1267	2	T21340
20	323	7.1	551	2	T30806
21	240	5.3	962	2	D6186
22	238	5.3	976	2	T51137
23	237.5	5.2	925	2	T06128
24	236.5	5.2	941	2	T51135
25	221.5	4.9	953	2	E84732
26	220.5	4.9	1039	2	T45779
27	219	4.8	951	2	T51132
28	216.5	4.8	925	2	T51133
29	215.5	4.8	933	2	C96495

30	213	4.7	950	2	T51134	ionotropic glutama
31	213	4.7	975	2	A84550	probable ligand-ga
32	211	4.7	923	2	P84732	probable ligand-ga
33	210	4.6	912	2	T51131	ligand gated chann
34	210	4.6	921	2	T51136	ionotropic glutama
35	208.5	4.6	1099	2	T16283	hypothetical prote
36	205	4.5	960	2	JE0356	gamma-aminobutyric
37	175	3.9	1103	2	JC5581	guanylate cyclase
38	159.5	3.5	1108	2	T59385	guanylate cyclase
39	151.5	3.3	1108	2	B55915	guanylate cyclase
40	140	3.1	965	2	T51244	N-methyl-D-asparta
41	139.5	3.1	920	2	B84640	probable ligand-ga
42	135.5	3.0	1323	2	I78557	N-methyl-D-asparta
43	133.5	3.0	1356	1	C45219	N-methyl-D-asparta
44	130.5	2.9	464	2	S75362	hypothetical prote
45	130	2.9	938	2	A46612	N-methyl-D-asparta

ALIGNMENTS

RESULT 1

JC7683

taste receptor T1R3 - mouse

C:Species: Mus musculus (house mouse)

C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 30-Sep-2001

C:Accession: JC7683

R:Kilgawa, M.; Kusakabe, Y.; Miura, H.; Ninomiya, Y.; Hino, A.

Biochem. Biophys. Res. Commun. 283, 236-242, 2001

A:Title: Molecular genetic identification of a candidate receptor gene for sweet taste

A:Reference number: JC7683; MUID:21222875; PMID:11322794

A:Contents: tongue

A:Accession: JC7683

A:Molecule type: mRNA

A:Residues: 1-858 <KIT>

A:Cross-References: DDBJ:AB049994

A:Comment: This protein, a seven-transmembrane receptor, serves as a mediator of the

A:Gene: tlr3

A:Map position: 4

C:Keywords: transmembrane protein

Query Match 71.4%; Score 3229; DB 2; Length 858;

Best Local Similarity 72.5%; Pred. No. 1.5e+245;

Matches 621; Conservative 67; Mismatches 147; Indels 22; Gaps 4;

QY 5 AVLGSLMALHPGCTGAPICLSQQLRMKGDYVIGLPGAEDEAGLSRTPSSPVC 64

DB 5 ATMGSLAAFLIEIGMGASICLSQOFKAQGDYILGILPLGSTBEATLNQRTQPN 64

QY 65 FSSNGILMALAKMAVEEINNKSDLLPGLRLGYDLFDCSEFVAMKPSLMLARAS 124

DB 65 FSPVLGFLMAKMAVEEINNGSALLPGLRLGYDLFDCSEFVAMKPSLMLARAS 124

QY 125 IAYCNITQYQPRVLAIVIGPHSSSELAMVYGFSEFLMPQVSYGASMEILSARETFPSEF 184

DB 125 IAYCNITQYQPRVLAIVIGPHSSSELATIGKFFSFLMPQVSYGASMDLSRETFPSEF 184

QY 185 RTVPSDRVOLTAABELLQEFEGNMWVAALGSDDEYGRQGISFSAALAAANGICIAHEGLVP 244

DB 185 RTVPSDRVOLTAABELLQEFEGNMWVAALGSDDEYGRQGISFSAALAAANGICIAHEGLVP 244

QY 245 QHDTSGLQGLKVLVDYLRQVNSKVQVYVFAARAVTSLEFSIHGGLSPKWWASPSMWL 304

DB 245 QHDTSGLQGLKVLVDYLRQVNSKVQVYVFAARAVTSLEFSIHGGLSPKWWASPSMWL 304

QY 305 TSDLVYGLPMAQMGTVLGLFQAGQLHEFPQYVTKHLATADPAFCNALGREGQGLEED 364

DB 305 TSDLVYGLPMAQMGTVLGLFQAGQLHEFPQYVTKHLATADPAFCNALGREGQGLEED 364

QY 365 VVGQRCPCQDCITTLQNVASGL-----NHQTSVYAAYVSVAQALHNTLQCNASGCP 416

DB 365 VVGQRCPCQDCITTLQNVASGL-----NHQTSVYAAYVSVAQALHNTLQCNASGCP 416

364 VVGQRCPCQDCITTLQNVASGL-----NHQTSVYAAYVSVAQALHNTLQCNASGCP 416

Db 535 REPTLVSLVQVPPSSNCRCLAGTRKGLIEGPTCCFEEVCEPDGEYSDTASAKN 594
 QY 541 CGODEMSERSTRCFRRSRFLWAGEPAVALLLLSLALGVLAAALFLPHHRDSPVQ 600
 Db 595 CPDDFMSMENTSCIAKEIFELSTWEPFGIALTLFAVIGITLFAVLGVFKFRPTLYK 654
 QY 601 ASGPPLA---CEGLVLCGLVCL-SVLLFPGQSPARCLAOQPLSHPLTGLSTLFLQAA 656
 Db 655 ATNRELSTLLFLSLC-----CFSSSLFFIGEPDWTCLRPAPGISFVLICSLIVNTN 710
 QY 657 EIFV---ESELPLS-----WADRLSGCLRGPMANLVLLAMVEVALCTWYLVAPPEVYT 709
 Db 711 RVLVFEAKIPTSFHRKMWGLNQ-----FLVFLCTEMQIYICVIMLYTAPSSYR 762
 QY 710 DMHMLPTALVHCRTSRVSWFGLAHATNATLAFCLFGLTFLVRSOPGCYNRARGTLFAML 769
 Db 763 NOLEDELIFITTCHEGSLMALGFLIGYTCLLAALCFEFAFNRKLPENFENAKFTTFEML 822
 QY 770 AFTTWSEFVLNAVQVLRPAVOMGALLLCVIGILAFAHLPRCYLLMRQPLNTPE 827
 Db 823 IFFIYIWSIFAPVASTYCKFVSAVEVIALASFGLLACIFENKLYIILFRPSRTIE 880

RESULT 5

159362
 calcium/polyvalent cation-sensing receptor precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 02-Jul-1996 #sequence-revision 02-Jul-1996 #text-change 05-Nov-1999
 C:Accession: I59362; A55594
 R:Ruat, M.; Molliver, M.E.; Snowman, A.M.; Snyder, S.H.
 Proc. Natl. Acad. Sci. U.S.A. 92, 3161-3165, 1995
 A:Title: Calcium sensing receptor: molecular cloning in rat and localization to nerve terminal
 A:Reference number: I59362; MUID:95241465; PMID:7724534
 A:Accession: I59362
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: mRNA
 A:Residues: 1-1079 <RES>
 A:Cross-references: EMBL:U02089; NID:9790578; PIDN:AAC52195.1; PID:9790579
 A:Experimental source: striatal
 R:Riccardi, D.; Park, J.; Lee, W.S.; Gamba, G.; Brown, E.M.; Hebert, S.C.
 Proc. Natl. Acad. Sci. U.S.A. 92, 131-135, 1995
 A:Title: Cloning and functional expression of a rat kidney extracellular calcium/polyvalent cation-sensing receptor
 A:Reference number: A55594; MUID:95110508; PMID:7816802
 A:Accession: A55594
 A:Molecule type: mRNA
 A:Residues: 1-133, 'X', 135-1079 <RIC>
 A:Cross-references: GB:U10354
 A:Experimental source: kidney
 C:Keywords: calcium; glycoprotein; phosphoprotein; transmembrane protein
 F:1-20/Domain: signal sequence #status predicted <SIG>
 F:187-212/Region: hydrophobic
 F:613-635/Domain: transmembrane #status predicted <TM1>
 F:650-670/Domain: transmembrane #status predicted <TM2>
 F:683-700/Domain: transmembrane #status predicted <TM3>
 F:725-744/Domain: transmembrane #status predicted <TM4>
 F:770-790/Domain: transmembrane #status predicted <TM5>
 F:806-828/Domain: transmembrane #status predicted <TM6>
 F:841-860/Domain: transmembrane #status predicted <TM7>
 F:901-921/Domain: transmembrane #status predicted <TM8>
 F:926-946/Domain: transmembrane #status predicted <TM9>
 F:949-969/Domain: transmembrane #status predicted <TM10>
 F:974/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status predicted
 F:989/Binding site: phosphate (Ser) (covalent) (by protein kinase A) #status predicted

Query Match 23.3%; Score 1054; DB 2; Length 1079;
 Best Local Similarity 30.3%; Pred. No. 2e-74;

Matches 269; Conservative 163; Mismatches 363; Indels 92; Gaps 28;

QY 9 LSWALL-HPGTGAPLCSQLRKAGVYVLCGLPPLG---BAERAGLSRTRPPSPVCTR 64
 Db 8 LALLALAHSSAYGP---DQRAKKRGDILIGLEPIHGVAKAKQDDKLS--RPESVEICR 62
 QY 65 FSSNGLLMALMKMAVEEINKSKDLLPGLRGYDLFDGCSPPVAMKRSMLFLA--KAGS 122

Db 63 YNRFQFRLQAMIFALIEINSSPILLPNMTLGYRIPTDCTNTVSALAEATLSFAONKIDS 122
 QY 123 RDIAYCNYTOYOPRVLAIVIGPHSSSELAMVYGFKEFSFLMPQVSYGASMEILSARETPS 182
 Db 123 LNLDFECMSHITSTIAVATGSGYSTAVANLLGLFIYIQVAYASSRLSNKNOYS 182
 QY 183 FFRVPSDRVOLTAABELLOEGFMNWAALGSDDEYGRGUSLIFSALAAAGCITAHGL 242
 Db 183 FLRTIPNDEHOTAMADIIIEYFRMNWGTIAADDYGRGIEKFRFEAEEDICIDISEL 242
 QY 243 VPLPRADSRIGKYODVHQNOSQVYVLLFASVHAHALFNYSISSRLSPKYNVASEA 302
 Db 243 I---SOYSDDEEQVVEVIONSTAKYIVFSSGPDLEPLKEIVRNITGRIMLASEA 298
 QY 303 WLTSDLVWGLPGMAOMGTVLGFLQGAOLHEFPQTV-KTHLALATDPAFCALGER---- 357
 Db 299 WASSSLIMPEYFHVYGGTIGFGLKAGIIPFRFLQVHVRKSVHNFACAEPEEFPNC 358
 QY 358 --EGLF-----EDVVGRCPCQD-----CITLQNVSA-----GLNHQ--TF 391
 Db 359 HLOGAKGPLVDITFVRSHEGGRNLLNSYAFRPLCTGDEINISVETPYMDYELRISY 418
 QY 392 SYTAAYSVQAALNLTLC-----NAGCPADDPYKPKQLLENMNLTF-HVGLPL 442
 Db 419 NVYLAIVSIHALODITYCLPGRGLFTNGS--CADIKVEAMOVAKHLHLFTNMKGQV 477
 QY 443 RFDSSGNVMEYDLKLV---VMQGSVPRLDHVGREN-GSLTERL-----KIRWHTSDNQ 493
 Db 478 TFDECGDLYGVNYSIINMHLSPEDSIV-FKEVGYNYAKAGERLFINEXEILMSGFSR 536
 QY 494 KPVSRCSRCQCEGOVRR-VKGFHSCYDCVDCERAGSTRONPDIACTFCGODEMSPERST 552
 Db 537 VPFNSCSHDCAGRGKGIIEBEPICCEVCPCPGEYSGEFDAACCKCPDPFMSNEMT 596
 QY 553 RCFRRSRFLWAGEPAVALLLLSLALGVLAAALFLPHHRDSPVQASGPPLA---CF 609
 Db 597 SCIAKEIFELMTWEPFGIALTLFAVLGIFLAFVIGVIRKRNPIYKATRELSTYLLF 656
 QY 610 GLVCLGLVCL-SVLLFPGQSPARCLAOQPLSHPLTGLSTLFLQAAEIFV---ESELPL 666
 Db 657 SLIC-----CFSSSLFFIGEPDWTCLRPAPGISFVLICSLIVKTRVLVFEAKIPT 712
 QY 667 S-----WADRLSGCLRGPMANLVLLAMVEVALCTWYLVAPPEVYTDHMLPTAL-V 720
 Db 713 SFHRKMWGLNQ-----FLVFLCTEMQIICIIWLYTPAPSSYRN-HELDEIIFI 763
 QY 721 HCRTRSWVSGFLAHATNATLAFCLFGLTFLVRSOPGCYNRARGTLFAMLAVITWSEFV 780
 Db 764 TCHESGLMALGSLIGYTCLLAALCFEFAFNRKLPENFENAKFTTFSLIFFIYIWSIFP 823
 QY 781 LLNAVQVLRPAVOMGALLLCVIGILAFAHLPRCYLLMRQPLNTPE 827
 Db 824 AYASYGKFAVSAVEVIALASFGLLACIFENKLYIILFRPSRTIE 870

RESULT 6

JH0561
 metabotropic glutamate receptor 2 precursor - rat
 C:Species: Rattus norvegicus (Norway rat)
 C:Date: 30-Jun-1992 #sequence-revision 30-Jun-1992 #text-change 13-Sep-1998
 C:Accession: JH0561
 R:Tanabe, Y.; Masu, M.; Ishii, T.; Shigemoto, R.; Nakanishi, S.
 Neuron 8, 169-179, 1992
 A:Title: A family of metabotropic glutamate receptors
 A:Reference number: JH0561; MUID:92110002; PMID:1309649
 A:Accession: JH0561
 A:Molecule type: mRNA
 A:Residues: 1-872 <TMAN>
 A:Experimental source: brain
 C:Comment: This protein is coupled to a G protein and evokes a variety of functions b
 C:Superfamily: metabotropic glutamate receptor 4
 C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane p
 F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-872/Product: metabotropic glutamate receptor 2 #status predicted <MET>
 F:568-590/Domain: transmembrane #status predicted <TRI>
 F:605-625/Domain: transmembrane #status predicted <TRI>
 F:637-655/Domain: transmembrane #status predicted <TRI>
 F:680-700/Domain: transmembrane #status predicted <TRI>
 F:726-747/Domain: transmembrane #status predicted <TRI>
 F:761-782/Domain: transmembrane #status predicted <TRI>
 F:795-819/Domain: transmembrane #status predicted <TRI>
 F:803-827/Domain: transmembrane #status predicted <TRI>
 F:832/Binding site: phosphate (Thr) (covalent) #status predicted

Query Match 14.8% Score 670.5; DB 2; Length 872;
 Best Local Similarity 27.8%; Pred. No. 2.1e-44;
 Matches 256; Conservative 133; Mismatches 385; Indels 147; Gaps 37;

1 MGPVAVLISLWALHPGTGAPLCSQOLRMKGDVYLGFLPL-----GEAEAGLRSTR 56
 4 ILGFLAL-LLMGVAEGPA-----KVLTEGDLVGLGFLPVHQKGPAGECGPVNEHR 57
 57 PSSPCTRESSGLMALAMKMAVEINNKSDLLPGLRGYDLFTCSEPVYAMKPSLMF 116
 58 -----GIQLEMLFALDRINRDPHLLPGVRLGAIHLDSCKDTHALQALDF 105
 117 LAKA-----GSRDI---AAYCNYTOYQPRVLAVIGPHSEELAMVYGRFFFLMPQVSY 167
 106 VRASLSRAGDSRHICPPGVSATHSDAPTAVYGVIGSGSDVSIQVANIILRLFOIPQISY 165
 168 GASMLLSARETFPSFFFTVPSDRVQLTAAELLQEPGMNVAALGSDDEYRGQGLSIFS 227
 166 ASTSAKLSKSRDYFARTVPPDFQAKAMAEILRFNMTYVSTVASSEGDYGETGLEAVE 225
 228 ALAAAGICIAHEGLVPLPRADSRIGKVQDYLHOVNOSVOVLLFASVHAALFNFS 287
 226 LEARKANICVAISEVGRAMSRAPAEVYRAIL---QKSAVAALFTRESEDAAREL--A 280
 288 ISSRLSPK-VWVASEAWLTSLDLYMGLPGMAQMGTVLGFLOGAQLHEFP--QYVKTAL 344
 281 ATQRLNASTVYVADSGWGLSEVVASGERAAGAIT-----IELASYISDFASYFQSL 334
 345 -----ATDPAFCGSAIGERQGLEEDYVVGORPCQCCITLQNTSAGLNHHQTSVTAAYS 399
 335 DPMNNSRNPMF-----REFMERFHCSPFQDCAHSHSLRAVPFQESKIME-VVNAVYA 387
 400 VQAALHN---TLQCNASG-CPAQDPVKPMQLEEN-MYNTLTFHVGGLP-----LAPDSG 448
 368 MAHALAHNRALCPNTHLCQMRPVGRRILTKDFVLNKKFPAAPRPADTDEVEFRDRFG 447
 449 NVDMEDYDLKLVWQGSVP-RLHDVGRFNGSLTERLTKIRWHT-SDNOKPVSRCSQCOBG 506
 448 DQIGRYNIFTYLRAGSGRRYKQVYMAEGLLDTSFLPMASPSAGPLPASCSCEPCLQN 507
 507 OVRAYKGFHSCCYDCVDCBAGSYRONPDIACTFCGQDEWSPERSTRCFRRRSRLAMGE 566
 508 EKVSVQPEVCCWICIPQOPYEYRL--DEFTCADCGILGWPNASITGCELFQEVYIRMD 565
 567 PAVILLILLSLALGLVLAAGLGFVHHRDSPLVOASGGGLAFCGLVCLVCL---SVLL 623
 566 AMAVPVITACIGALATLFLVGLGFVRRHNAITPVYKASGRGL-CY--TLGGLVCLVCLV 622
 624 FPGQSPARCLAOQPLSLHPLTGLSTFLQA---AEIF-----VESELPSMADRISG 674
 623 FLAKSTAVCTLRIGLGTAFSVCSALLTKTNRIARIFGARREGAQRFRFISPSQVAI 682
 675 CLRGFWAMLVLLAMLEVALCTWYLVAF-----PEVTDWMHLPTLEALVHCRT 724
 683 C-----LATISQOLLIVA--AMLVVEAPGTGKETAPERREVV-----LRCHN 723
 725 RSWVSPGLAHNNTALPLCLGTFLVVSQPCYNRARRITLTAALAYFITWVSFPLAN 784
 724 KASMLG-SLATNVLLIALCLTLYAKTRCPENFENAKFISFTMTYTCITLWALFPIF-- 780
 785 VOVLPRAPVOMGALLC-----VLGIILAAFLPRCYLLMRQGLN-----TPEF 828

DB 781 --YVSSDVRVQVTTMCVSVSLSGSVLGLCLPA-----PKHILTRPQKRVSHRAPTSRF 835
 829 FLGGPBGDAQGGNDGNTGNQG 849
 DB 836 -----GSAAPRASNLG-QG 849

RESULT 7

A:6742 metabotropic glutamate receptor, mglur6 - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 21-Sep-1993 #sequence_revision 18-Nov-1994 #text_change 20-Jun-2000

C:Accession: A46742

R:Nakajima, Y.; Iwakabe, H.; Akazawa, C.; Nawa, H.; Shigemoto, R.; Mizuno, N.; Nakani

J. Biol. Chem. 268, 11868-11873, 1993

A:Title: Molecular characterization of a novel retinal metabotropic glutamate recepto

A:Reference number: A46742; PMID:93280152; PMID:8389366

A:Accession: A46742

A:Status: preliminary

A:Molecule type: nucleic acid

A:Residues: 1-871 <NAK>

A:Cross-references: GB:D13963; NID:9391856; PIDN:BA03066.1; PID:9391857

A:Experimental source: retina

A>Note: sequence extracted from NCBI backbone (NCBIN:133246, NCBIP:133250)

C:Superfamily: metabotropic glutamate receptor 4

C:Keywords: G protein-coupled receptor; transmembrane protein

Query Match 14.0% Score 631.5; DB 2; Length 871;
 Best Local Similarity 25.6%; Pred. No. 2.5e-41;
 Matches 225; Conservative 146; Mismatches 410; Indels 99; Gaps 31;

4 PAVLGISLWALHPGTGAPLCSQOLRMKGDVYLGFLPLGAEAEAGLRSTRSPVCT 63
 5 PVLMLWALWISQAGIA---CGAGSVRLAGLGLTGLGFLPVHARGAAG-----RAGC 52
 64 RF-SSNGLLMALAMKMAVEINNKSDLLPGLRGYDLFTCSEPVYAMKPSIME-----LA 118
 53 ALKKEQGVRLRLMALXALDRVNAADPELLPGVRLGAILDTCRDIYALQALSFQALIR 112
 119 KAGSRDIAVYNYTOYQPR-----RVLAIVIGPHSEELAMVYGRFFFLMPQVSYGASM 171
 113 GKGDDDEASVRCPPGPPRLSAPPERVAVVAGASASVIMAVNYLRLAIPQISVASTA 172
 172 ELLSARETFPSFFFTVPSDRVQLTAAELLQEPGMNVAALGSDDEYRGQGLSIFSALA- 230
 173 PELSDSTRIDFESRVVPPDSYQAQAVDIVRALGMNYSVTLASSEGYGEAFVQISR 232
 231 AARGICIAHEGLVPLPRADSRIGKVQDYLHOVNOS-SVOVLLFASVHAALFNFSIS 289
 233 EAGVCYCIADS--IKTIPR--EPKPEGHKVIIRLMETPNARGLITIRANEDDIRVLEATRQ 288
 290 SRLSPK-VWVASEAWLTSLDLYMGLPGMAQMGTVLGFLOGAQLHEFPQYVTH----- 341
 289 ANLTGHFLWVGSNMGSKISPLINLEERA-VGAIT-ILPKRASIDQFDQYFMTSLNNR 346
 342 ----LALATDPAF-----GALGERQGLE-----EDVVGORPCQCCITLQNTSAGLNHHQ 389
 347 RNIMFAEWEENFNCKLSSGQSDSTRKCTGEEITGQD-----SAVEQEGK 394
 390 TFSYAAVSVQAALHNTLQCNASG---CPAQDPVKPMQLEENMYNLTFH-VGGLPLRF 444
 395 VQFVIDAYAIHALHSHQALCPGHTGICPAMEPIDGTLLHYIRAVFNKSAGTPVWF 454
 445 DSGNVDMEYDL-KLVWQGSVPR--RLHDVGRFNGSLTERLTKIRWHTSDNOKPVSRCSR 501
 455 NENGAPGRYDIFQYQATNGSASSGQYQAVGWMAELRLDMDEVLRMSGDPHEVPQSCGL 514
 502 QCOEGQVRR-VGFHSCCYDCVDCBAGSYRONPDIACTFCGQDEWSPERSTRCFRRRSR 560
 515 PCGPBEKRMVAGV-PCMCHEACD--GYRFQYDFTCEACGDMRPPNHHGCRPTPV 571
 561 FLAMEPVALILLILLSLALGLVLAAGLGFVHHRDSPLVOASGGGLAFCGLVCLVCL 620

```

Db 572 RLWSSPPAALPLLLAVGIMATTTTMAFEMRHNDPTIVRASGRELSTVLLTGIFLIYAI 631
QY 621 VLLPGOPSPARCLAOPLSHLPLTGLSTLFLQA---AEFVSESLPLSWADRLSGCLR 677
Db 632 TELWAVECAICARRLGLGTTLSALTTRKTRIRYRFEQGRKRVTPPTPLIS----- 687
QY 678 GPWMLVYL--LAMLVEVALCTWYLVAFPEPVTDHMLPT-----EALVHCRTSNVWS 729
Db 688 -PTSQVLTFFGLTSLQVGVYIAM--LGAOPPHSVIDYEOQFVDEPQAGVLCQMSDLSL 745
QY 730 FGLAHATNATLAFICFLGTFLVRSQPGCYNRRARGLTFEMLAFYTWVSFVLLANV---- 785
Db 746 IGCL-GTSLMLMTCTYTAIRARGVPEFENAKRPIGFMTTCITLWLFVIFFGTAQSA 804
QY 786 -QVVLRAVOMGALLCYLILAAFHLPFCYILMRQPLN 824
Db 805 EKIIQTTLTVSLTSLASVSLGMLVYVKTYVILFHPEDN 844

RESULT 8
JH0563
C:Metabotropic glutamate receptor 4 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 17-Nov-2000
C:Accession: JH0563; I58149
R:Tanabe, Y.; Masu, M.; Ishii, T.; Shigemoto, R.; Nakanishi, S.
Neuron 8, 169-179, 1992
A:Title: A family of metabotropic glutamate receptors.
A:Reference number: JH0561; M01D:92110002; PMID:1309649
A:Accession: JH0563
A:Molecule type: mRNA
A:Residues: 1-912 <TAN>
A:Experimental source: brain
R:O'Hara, P.J.; Sheppard, P.O.; Thøgersen, H.; Venezia, D.; Haldeman, B.A.; McGrane, V.;
Neuron 11, 41-52, 1993
A:Title: The ligand-binding domain in metabotropic glutamate receptors is related to bac
A:Reference number: I58149; M01D:9333699; PMID:8338667
A:Accession: I58149
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-123; 'R', 125-912 <RES>
A:Cross-references: GB:M90518; NID:g205400; PIDN:AA93190.1; PID:g205401
C:Comment: This protein is coupled to a G protein and evokes a variety of functions by m
C:Genetics:
A:Gene: GRIIR4
C:Superfamily: metabotropic glutamate receptor 4
C:Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; phospho
F:1-32/Domain: signal sequence #status predicted <SIG>
F:33-912/Product: metabotropic glutamate receptor 4 #status predicted <MET>
F:588-610/Domain: transmembrane #status predicted <TRI>
F:625-645/Domain: transmembrane #status predicted <TII>
F:657-675/Domain: transmembrane #status predicted <III>
F:700-720/Domain: transmembrane #status predicted <TRV>
F:751-772/Domain: transmembrane #status predicted <TRV>
F:786-807/Domain: transmembrane #status predicted <TVI>
F:822-847/Domain: transmembrane #status predicted <VII>
F:98,301,454,484,569/Binding site: carboxylate (Asn) (covalent) #status predicted
F:621,689,695,859,870/Binding site: phosphate (Ser) (covalent) #status predicted

Query Match 13.7%; Score 619.5; DB 2; Length 912;
Best Local Similarity 24.6%; Pred. No. 2.3e-40;
Matches 222; Conservative 145; Mismatches 397; Indels 137; Gaps 31;

```

```

Db 123 VOALIEKDGIEVRGSGGPPITTRPERVGVIGASGSVSMVANIILRFKIPGISYST 182
QY 171 MELLASREFPSPFRYPSDRVOLTAALAELOEGRMMVVALGSDDEYGRGLIFSAIA 230
Db 183 APDISDSNRDFFFRVPSPTYQANAVDIYALKMNTVSTLASEGSGEYVAFYOKS 242
QY 231 AAR-GICIAHEGLVLPRAADSRIGKYQDVLIHQ--VNQSSVOVLLFAS-----VHAH 281
Db 243 RENGVACIAQS--VKIPR--EPKTEGDKIKRILETSMNAGIIIFANEDDIRVLEAR 298
QY 282 ALFKYSSIRLSPPKVAASEAM--LTSDLVGLPQMAOMGYLGLQGAOLHEPQVYKT 340
Db 299 R-----ANQGFHEFWMGSDWSGSKSAPVLTLEVAE--GAAT--ILPRMSTRGDRYSS 350
QY 341 H-----LALADPAPFCSALGEREGLDEEDVVGORPODCITLQNVASGLNHHQ 389
Db 351 RTLDNNRNITWFAFEWEDNFHCKL---SRNALKKGSHKKCTNERNIGOD--SAVEDEK 405
QY 390 TFSYAAVYSVAOALHNTLQCNASG---CPADPVKPMQILLENMNLTEH--VGLPLRF 444
Db 406 VQFVIDAVYAMGHALHAMHBDLCPRGVGLCPRMDPVDGTQLKXIRVNFSGIAGNPVTF 465
QY 445 DSGNVDMEXYDLKLMWQGSVPRLDVGRNGSLRTERLKRIRHTSDNOKPVSCRCQO 504
Db 466 NENGDAPEGRTDIYQYQLKNGSAEYKVGSWTDHLRLIERMOWGSGQOLPRISICLPQ 525
QY 505 EGQVRR--VKGFHSCCYDCVCEAGSYRONPDDIACFCGODEMSPERSTRFRRSFELA 563
Db 526 PGERKATYKGM--ACQWNCERPT--GYQYQDRTYCKKCPIDMRTENRTSQPIPIYKLE 582
QY 564 WGEPAVLLLLLSLALGLVLAALGLFVHNRDPLVOASGGPLACFGLVGLVCLSVLL 623
Db 583 WDSFWAVLPLEFLAVYAGIATLFLVVTFVRYNDPFIYKASGRELSTVLLAGIFLCYATFTL 642
QY 624 FPGQSPARCLAOPLSHLPLTGLSTLFLQAALF-----VSESLPLSWADRLSG 674
Db 643 MIAPBDLQTCRLRIEYGLGMSISYALLTITNRIYRFEQGRKRSVASPRETISASOLA 701
QY 675 CLRGPMALVYLILMLVEVALCTWYLVAFPEPVTDHMLPTLALVCRTRSNVWSFGLAH 734
Db 702 -----IFILISLQDLGLGVAFVVD--PESHVVDQDQDTLDRPAR-----GVLK 745
QY 735 ATNATLAFLCFLGFLV-----NSQPCYRNRARGLTFEMLAFYTWVSFVPL 782
Db 746 CDISDSLICLLGYSMLMTCTYTAIRARGVPEFENAKRPIGFMTTCITWLAFLIF 805
QY 783 -----ANVOVLRPAVOMGALLCYLILAAFHLPFCYILMRQPLNTP 826
Db 806 FGTSQADKLIYQTTTLTVSVLSASVSLGML-----YMPKYIILFHPEDNVP 854
QY 827 E 827
Db 855 K 855

RESULT 9
S71376
C:glutamate receptor homolog - cherry salmon
C:Species: Oncorhynchus masou (cherry salmon)
C>Date: 11-Mar-1998 #sequence_revision 17-Apr-1998 #text_change 01-May-1998
C:Accession: S71376
R:Kubokawa, K.; Miyashita, T.; Nagasawa, H.; Kubo, Y.
FEBS Lett. 392, 71-76, 1996
A:Title: Cloning and characterization of a bifunctional metabotropic receptor activat
A:Reference number: S71376; M01D:96354880; PMID:8769318
A:Accession: S71376
A:Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 1-1218 <KUB>
C:Keywords: glycoprotein; phosphoprotein
F:603-625/Domain: transmembrane #status predicted <TM1>
F:640-660/Domain: transmembrane #status predicted <TM2>

```

F:672-690/Domain: transmembrane #status predicted <TM3>
 F:717-737/Domain: transmembrane #status predicted <TM4>
 F:761-782/Domain: transmembrane #status predicted <TM5>
 F:786-817/Domain: transmembrane #status predicted <TM6>
 F:826-850/Domain: transmembrane #status predicted <TM7>
 F:104,233,403,525,757/Binding site: carbohydrate (asn) (covalent) #status predicted
 F:636,699,961/Binding site: phosphate (Ser) (covalent) (by protein kinase C) #status pre
 F:705/Binding site: phosphate (Thr) (covalent) (by protein kinase C) #status predicted
 F:892/Binding site: phosphate (Ser) (covalent) (by CAM-dependent kinase) #status predic

Query Match 13.5%; Score 612.5; DB 2; Length 1218;
 Best Local Similarity 25.7%; Pred. No. 1.2e-39;

Matches 226; Conservative 140; Mismatches 370; Indels 143; Gaps 34;

QY 30 RMKGIVYVGLFPL-----GEAEAGLRSTRSPSSVCTRSSNGILMFLA 75
 DB 46 RMDGDIILGALFVHHQDSTQVARKGGEVRE-----OYGIQRYEA 87
 QY 76 MKMAVEINNKSDLLPGRLGYDLFDTCSEPVVAMKPSIMFLA-----KAGSRD 124
 DB 88 MFHLIDRINSNPILLPTTLCCEIRDSCHSSVALLQSEIRDSLSIRDDDKDGTSR 147
 QY 125 IAAVONTYOYR-----RVLAIVGRHSELAMVTGKFFSFLMPQVSYGASMEILLARET 179
 DB 148 QMCIEGMPPSQPAPTKRPIAGVIGSGSSVAIQVONLLQLFNIPQIAYSATSIDSLDKTL 207
 QY 180 FPFSEFRVPSRVOITAAELLQEGWVVAALSDDEYRGQGLSIFALAAKICITAH 239
 DB 208 FKYLKRVPSDTLQARALDLVKRNMVTVSAVHTGEGSGMAFELASQEBELCTAH 267
 QY 240 EGLVPLPRAD---DSRLGKVDVLHQVQSSVQVLLFASVHAHALNFYSISRLSPKV 296
 DB 268 SDKIYSNKGKHFRLRLKLRRL-----PKARVVVCECEMTVAGLL---MAMR---RL 316
 QY 297 WVASEAWLTSLDWGLPQMAQGYVL-GFLQRG-----AQDH---EPFYR-VKTHLAL 344
 DB 317 GVAGGF-----LLIGSDGMADRDEVEGEYEQEAVGIVTKLHSEVTSFDFYFLKRLNLT 371
 QY 345 AT-PPAFCSALGERRQ-----GLEEDVVGQRCPQDCIT---LQNVASGLNHHTFESVY 394
 DB 372 NTRNPFEPFQOHROCFIRGHLENNMYKKNKSGYESLEDNYVDSDMGF-----VI 424
 QY 395 AAVYSVAQALHNTLQCNASG---CPAODPYKPMQLLENMTNLTP-HVGGPLRLPDESSGN 449
 DB 425 NAITAMAGGLDMHSLCPGHVGLCKAMPIDGSQLLEFLMRTSTGVYSGEDVWDEMGD 484
 QY 450 VDMETDLK--LAWOGSVPLRHLDVGRFNGSLRTERLKITRMHTSDNQKPYSCRSQCQEQ 507
 DB 485 TPGRYEIMNLQYVERGADYINVGSWHBGQLSIDDYMQINRSD--MVLVSCSEPCSKGE 542
 QY 508 VBRV-KGFHSCCYDVCDEAGSYRQNPDDIACFPQGDQWSPERSTRFRRRSRFLANGE 566
 DB 543 IKVIRKGGVSCWICTACKDNEIYO--DEFTCTACDGLGMMWDPDELEGEPTTLRLLEWGN 600
 QY 567 PAVLLILLLSLAIGLVLAALGLFVHNHDSPLVQASGGLACFGL---VCLGLVLCVYL 623
 DB 601 PESIVQVFAELGILVTSFVTFEIVLYRDPYVSSREL-CYLLAIFLGIOTPEPLI 659
 QY 624 FPGQSPARCIAQOPLSHLPLTGLSTLFLQAAELFEVSELPLSWADRLSG---CLRG 678
 DB 660 --AAPTVAASYCLRLVLGLSATMWCYSALVTKTNR-----ARILASKKKIKCTRK 707
 QY 679 P-----NMLIVVL-LAMLYEVALCTWYLVAFPRPEVYTDMMHLPTALVHCTRSMVSGL 732
 DB 708 PRFSAMQVLIAGLVSVQLLEVTLLILBPMVPKSYPSI-REVFLICMT-STVGMYA 765
 QY 733 AAHTNATLAFCLGTFLVRSQPCYNARAGLTFAMLAFTTWSEFVLL--ANQOYLR 790
 DB 766 PLGNGLLIMCTYYAFTRVNPANFNKAKIATFTMTTCIIMLAFVVIYIGSNKII-- 823
 QY 791 PAVOMGALLLCVGLIILAAFLPRCYLLMRQPLNTPPEFF 829
 DB 824 --TTSFVSLSVTVLALGCMFSPKIYIILAKPERNVSFAF 860

RESULT 10 A49874

metabotropic glutamate receptor 7 - rat

N:Alternate names: metabotropic glutamate receptor mglur7

C:Species: Rattus norvegicus (Norway rat)

C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999

C:Accession: A49874; I57954

R:Okamoto, N.; Hori, S.; Akazawa, C.; Hayashi, Y.; Shigemoto, R.; Mizuno, N.; Nakanis

J. Biol. Chem. 269, 1231-1236, 1994

A:Title: Molecular characterization of a new metabotropic glutamate receptor mglur7 c

A:Reference number: A49874; MUID:94117433; PMID:8288585

A:Accession: A49874

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-915 <RES>

A:Cross-references: GB:D16817; NID:9458728; PIDN:BA04092.1; PID:9458729

R:Saugstad, J.A.; Kinzie, J.M.; Mulvihill, E.R.; Segerson, T.P.; Westbrook, G.L.

Mol. Pharmacol. 45, 367-372, 1994

A:Title: Cloning and expression of a new member of the L-2-amino-4-phosphonobutyric a

A:Reference number: I57954; MUID:94195260; PMID:8145723

A:Accession: I57954

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-915 <RES>

A:Cross-references: EMBL:U06832; NID:9459657; PIDN:AAA20655.1; PID:9459658

C:Gene: MGLUR7

C:Superfamily: metabotropic glutamate receptor 4

C:Keywords: neurotransmitter receptor

Query Match 13.4%; Score 604.5; DB 2; Length 915;
 Best Local Similarity 23.9%; Pred. No. 3.5e-39;
 Matches 212; Conservative 169; Mismatches 386; Indels 121; Gaps 32;

QY 6 VIGSLMALLHNGCTAPCLSQLRMKGIVYVGLFPGLEAEAGLRSTRSPSSVCTRF 65
 DB 20 VLEVLTCLVLAARQEMTAHPSIKIRIBGDTYLGGLFPYHAKGPSV-----FCGDIK 71
 QY 66 SSGILMALAMKMAVEINNKSDLLPGRLGYDLFDTCSEPVVAMKPSIMFLAKAGSRDI 125
 DB 72 RENGTHREAMLYALDQINSDPNLLPNTVIGARIDFTCSRDYVALLQESTFEVALIQKDT 131
 QY 126 A-AVCNTYOYR-----RVLAIVGRHSELAMVTGKFFSFLMPQVSYGASMEILLAR 177
 DB 132 SDVRC--TNGEPVVFKEKGVVIGAGSSVSIVANILRLFLQIPDISYASTABELSD 189
 QY 178 EFPFSEFRVPSDRVQUTAAAEELLQEGWVVAALSDDEYRGQGLSIFALAAKICITAH 236
 DB 190 RYDFPSKRVPPDSQADAMDIYKALGMVYVSTLASGSGYGEKVESTQISKRAGLIC 249
 QY 237 IAHGELVPLPRADSRIGKVDVLHQ-VNQSSVQVLLFAS-VHAHALNFYSISRLSP 294
 DB 250 IAQSVRIQERKD--RTIDFRIIKQLDTPRSRAVVFANDEDIKQILAAKRAYQVGH 307
 QY 295 KVVWASEAWLT-----SLVNG-LPQMAQGYVLF-----LQKQGLHEFYQY 337
 DB 308 FLWVGSDSWGSKINPLHGHEDIAEGALITQPRRAVVEGFDAFTSRTLENNRNWFAEY 367
 QY 338 VKTHLALTDPAFCALGERBQGLEEDVVGQRCPCDCITLQNVASGLNHHTFS--YU 394
 DB 368 WEENFNC-----LTSSSKEDIDRKCTGQ-----ERIKDSNYDEGVQYVI 412
 QY 395 AAVYSVAQALHN--TLQCNASG--CPAODPYKPMQLLENMTNLTPH-VGGLPLRDESSGN 449
 DB 413 DAVYAMAMALHNMKNDLQADRYGVCPEMEQAGKKLTKYIRIVNNGSAGRPVWENKNGD 472
 QY 450 VDMETDLKAWQGSV-PLRHLDVGRFNGSLRTERLKITRMHTSDNQKPYSCRSQCQEQ 508
 DB 473 APGRYDIPOYQTTNTNGYRLIGQWTDLQINTEDMQMGVREIPSSVCTLLPCKPQOR 532
 QY 509 RRVGFHSCCYDVCDEAGSYRQNPDDIACFPQGDQWSPERSTRCFRRSRFLAWGBPA 568

```

Db 533 KTKQKGPCCWTCPCPCD--GVOYQFDEMTQCHQCPYDPRNENRNGCONIPITIKLEHSPW 590
Oy 569 VILLLLLLSTALGLVLAALGLFVHHRSPLVOASGGLAGLGLVGLVGLVGLVGLVGLVGLV 628
Db 591 AVIVFLAMGIATFIVMATFIRKNDPIYRASGRELSTVLLGLVGLVGLVGLVGLVGLVGLV 650
Oy 629 SPARCLAOQPLSHLPGLGCLSTLFL-----QAEIIVESELEPLSMADRLSGCLGPMWML 683
Db 651 DVAVCSRRFVE--LGLMGCSYALLTKTNRIYRIFEGKKSVT--ADRLI-----SPISQL 703
Oy 684 VLLAMVEVALC---TWLYAPPEVYTDH---MLPTEALVHCRTKRWVSFGLAHAT 736
Db 704 AI--TSSLISVQLGVEITWFGVD--PPNITIIDKHTNNPQAR-----GVKCD 750
Oy 737 NATLAFCFGLTFV-----RSQGCYNRAGLGFALAYFITWVSFVPL--LA 783
Db 751 ITDQITISLGLSTILMATVCAIKTRGVPEENNEAKPIGFTWYTCIYVLAIFIPFG 810
Oy 784 NVQVYLRPAVOMGALLCY-----LGIILAFHLPRCYLLMRPGIN 824
Db 811 TAGAERKIYQITLTITISMNLASVALGML---YMPKYIITIFHELN 855

```

RESULT 11

```

metabotropic glutamate receptor 3 precursor - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 13-Sep-1998
C:Accession: JH0562
R:Tanabe, Y.; Masu, M.; Ishii, T.; Shigemoto, R.; Nakanishi, S.
Neuron 8, 169-179, 1992
A:Title: A family of metabotropic glutamate receptors.
A:Reference number: JH0561; M0ID:92110002; PMID:1309649
A:Accession: JH0562
A:Molecule type: mRNA
A:Residues: 1-879 <TAN>
A:Experimental source: brain
C:Comment: This protein is coupled to a G protein and evokes a variety of functions by
C:Superfamily: metabotropic glutamate receptor 4
C:Keywords: G protein-coupled receptor; glycoprotein; phosphoprotein; transmembrane prot
F:1-22/Domain: signal sequence #status predicted <SIG>
F:23-879/Product: metabotropic glutamate receptor 3 #status predicted <MET>
F:577-599/Domain: transmembrane #status predicted <TRI>
F:614-634/Domain: transmembrane #status predicted <TRI>
F:646-664/Domain: transmembrane #status predicted <TRI>
F:689-709/Domain: transmembrane #status predicted <TRI>
F:735-756/Domain: transmembrane #status predicted <TRI>
F:770-791/Domain: transmembrane #status predicted <TRI>
F:804-828/Domain: transmembrane #status predicted <TRI>
F:209,292,414,439/Binding site: carbohydrate (asn) (covalent) #status predicted
F:610,845/Binding site: phosphate (Ser) (covalent) #status predicted

```

```

Query Match 13.28; Score 597.5; Db 2; Length 879;
Best Local Similarity 25.28; Pred. No. 1.2e-38;
Matches 218; Conservative 142; Mismatches 380; Indels 125; Gaps 32;

```

```

Oy 25 LSQQLRMKGDVYLGLPLGE---AAEAGLRSTRSPVCTFRSSNGLLMALAMKAV 80
Db 29 MRRRIKIEGLVGLGFLPINKGTGTECGRINER-----GIORLEAMLFAI 76
Oy 81 EFINNKSDLLPGLRLGYDLPDTCSEPVYAMKPSIMPL--AKAGSRDIAY--CNTYOQ-- 135
Db 77 DEIKNDVYLLPGVXGLVHIDTCSRDTYALEQSLFVRASLTKVDEAYMCPDGSVAIQE 136
Oy 136 --PRVLA-VIGPHSSSLAMTGRKFFFLMPQVSYGASMEILLSAREPPSPFRVPDGRV 192
Db 137 NIPLLINGVIGSYSSVIOVANLRLFLPQISYASTSAKLSKSRDYFARTVPDPFY 196
Oy 193 QLTAAAELEDFGNWVAALGSDDEYGRGSLISALAAARGCITAHGVLPLPRADSR 252
Db 197 QAKMAELIRFNTYIVSYVASEDYGTEIGLEAFQEARLRNLCITATAEKYGSRNIRKSY 256

```

```

Oy 253 LGRVQDYLHOVNOSSVOVLLFASVHAHALFNYSISSRLSPK--VWVASEAW--LTSOLV 310
Db 257 DSVIRRELLQKPN---ARVVLFMKSDSRELI--AAANRVNASFTVVASDGMGAQESIVK 311
Oy 311 GLPRMAQMGIVLGLQAGADLHPQYVKI-----HLATLPDAF-----CSALG 355
Db 312 GSEHVAAGATLELASH--PVRODFRFOSLINPNYNN-----RNWFEDFEQKFOCSLON 365
Oy 356 EREGLEDVEDVQRCPCDCITLQNVASAGLNHHQTFESVYAAYSVAQALH--NTLQCN 412
Db 366 KRNH-----ROYCDHLALIDSNTYE---QESIMVYVNAVYAMAHALKMORTLCPNT 415
Oy 413 SG--CPAADPYKPMQL--ENMYNLTFHVGLP-----LRPSSGNDMEYDLKLMWOG 463
Db 416 TKLQDAMKIIDGKKLYKEYLIKINFAPFNPKNKADSIKVEFDGDMGRYNNFNLOOTG 475
Oy 464 SVPRLDHVGFRNGLSRERLKIEMHTSDNCKPVSRGROEQEGRVRKGFHSCCYCDV 523
Db 476 GKYSYLKVGHMAETLSLDVDSIHW--SRNSVPSQSDPCAPNEMKMOGPDVCCWICIP 533
Oy 524 CEAGSYRQNPDDIACFCGODEMSPERSTRCFRRSRFLAMGEPAVLLLLLSTALGLV 583
Db 534 CE--PYEIVDEFQCMDCGCGQWPTADLSGCYNLPEDYIMEDMAIGPTIACGLFCT 591
Oy 584 LALGLFVHHRSPLVOASGAPLACGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLV 643
Db 592 CIVTVFIRKNNPPLVASGRELICYILLFGVSLSYCTFFFIAPSPICALRGLGTS 651
Oy 644 LTGLSTLFLQAA--AEIIVESELPLSMADRLSGCLNG-----PMWLVYLLAM--LV 691
Db 652 FAITYSALLTKTNIAIF-----DGVKGAQRPKRTISSSOVFTGLGILV 698
Oy 692 EVALCTWYLAFFPEVYTDHMLPTE---ALVHCRTSRWVSFGIAHATNMTLAFELGT 748
Db 699 QIVWVSWLLETPG--TRRYTLEKREYVLIKCNVAD--SSMLSTLYDVVLVILCTVYA 755
Oy 749 FIVASQGCYNRAGLGFALAYFITWVSFVPLIANOVYLRPAVOMGALLC----- 801
Db 756 FKTRKCEPNEENAEKIFFTWYTCIYVLAIFP---YVTSDDYRVQTTMCTISVLSG 811
Oy 802 --VGIILAFHLPRCYLLMRPGIN 824
Db 812 FVVLGCLFA---PRVHLYLFOPOKN 833

```

RESULT 12

```

metabotropic glutamate receptor 8 - mouse
C:Species: Mus musculus (house mouse)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 24-Sep-1999
C:Accession: I49142
R:Duvoisin, R.M.; Zhang, C.; Ramonell, K.
J. Neurosci. 15, 3075-3083, 1995
A:Title: A novel metabotropic glutamate receptor expressed in the retina and olfactory
A:Reference number: I49142; M0ID:9529344; PMID:7722646
A:Accession: I49142
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-908 <RES>
A:Cross-references: EMBL:U17252; NID:g854728; PIDN:AAA68149.1; PID:g854729
C:Genetics:
C:Gene: mglur8
C:Superfamily: metabotropic glutamate receptor 4
C:Keywords: neurotransmitter receptor

```

```

Query Match 13.18; Score 593.5; Db 2; Length 908;
Best Local Similarity 23.48; Pred. No. 2.3e-38;
Matches 206; Conservative 157; Mismatches 368; Indels 149; Gaps 29;

```

```

Oy 26 SQQLRMKGDVYLGLPLGEAEGLRSRTSPVCTFRSSNGLLMALAMKAVEINN 85
Db 37 AHSIRLDGDIILGLFVYHAKGENGV-----PCGDLKREKGIHLEMLAIAIDQTNK 88

```



```

0Y 471 ---VGR--NSLTERKJIMHMSDNOKPVS--GROOCEGVRRV-KGFHSCTDYC 522
Db 480 DYIVGSMNDNELMKDDEEV-WSKSN---IIRVCSEPECKGQIKIRGEVSCWTCI 535
0Y 523 DCEAGSYRONNDIACJFCGODEMSPERSTCFRRRSRFLANGEPAVLILLLLSIALGI 582
Db 536 PCKENEYVF--DEYCKACKQJGSMPTDDLIGCDDILPQYLIRMDGPDEIAVVFACIGLLA 593
0Y 563 VIALAGLFVHRDSEFLVOASGGPLACGCL--VCLGLVCLSVLLFPGQSPARCLAQ-- 637
Db 594 TLEVVAVFIIIRDPVPVWSSSREL-CYIIAGICIGLCTFCFLI--AKPQIICYLQRI 650
0Y 638 -PLSLPLUTGCLSTFLQAAEIFESELPLSMADR--LSCLGPNAMLVLLVLLAMEYA 694
Db 651 IGLSPMSYSALVYTKTRINARIILGSKKKICTKPRMSSCAOLVIAFLILICIGLIVA 710
0Y 695 LCTWLVLPPEVVYTDWMLPTEALVLCRFRSWSFGLAHATNATLAFCLFGLTFVRSQ 754
Db 711 L-----FINEPDIHDHPDSINEVYLLICNTNMLGVVPELgV--NSLLISTFYAFKTRNV 764
0Y 755 PGCYRANGLFFAMLAFTITVWSVPLL--ANVOV-----LRPAVQGALLCYLGI 805
Db 765 PANNEEAYIAFTYTCIILAVCPYIFSNYKIITMCFSVSLASATVALGCMF----- 818
0Y 806 LAAFILPCYILMROPLNPEPF-----LGGPBGDAQONGNENTGNQK 850
Db 819 -----VKKVYIIILKPERNVSAFTTISTVVRMHAGDDKSSSAARSSSLVNLMK 867

```

Search completed: May 19, 2003, 09:51:18
Job time : 25.1657 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:46:50 ; Search time 11.0911 seconds
(without alignments)

3186.136 Million cell updates/sec

Title: US-09-927-315-15
Perfect score: 4524
Sequence: 1 MGPAYGLSLMALHPGTC.....GPDAGQGNQNTGNGCKHE 852

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 112892 seqs, 41476328 residues

Total number of hits satisfying chosen parameters: 112892

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_40:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1068.5	23.6	1078	1 CASR_HUMAN	P41180 homo sapien
2	1062	23.5	1079	1 CASR_MOUSE	Q99y96 mus musculu
3	1060	23.4	1085	1 CASR_BOVIN	P35384 bos taurus
4	1054	23.3	1079	1 CASR_RAT	P48442 rattus norv
5	670.5	14.8	872	1 MGR2_RAT	P31421 rattus norv
6	658.5	14.6	872	1 MGR2_HUMAN	Q14416 homo sapien
7	631.5	14.0	871	1 MGR6_RAT	P35349 rattus norv
8	622	13.7	912	1 MGR4_HUMAN	Q14833 homo sapien
9	619.5	13.6	877	1 MGR6_HUMAN	P31423 rattus norv
10	617	13.6	877	1 MGR6_HUMAN	Q15303 homo sapien
11	615.5	13.6	915	1 MGR7_HUMAN	Q14831 homo sapien
12	612.5	13.5	976	1 MGR_DROME	P91683 drosophila
13	604.5	13.4	877	1 MGR3_HUMAN	Q14833 homo sapien
14	604.5	13.4	915	1 MGR7_RAT	P35400 rattus norv
15	597.5	13.2	879	1 MGR3_RAT	P31422 rattus norv
16	594.5	13.1	908	1 MGR8_RAT	P70579 rattus norv
17	583.5	13.1	908	1 MGR8_MOUSE	P47743 mus musculu
18	585.5	12.9	908	1 MGR8_HUMAN	O00222 homo sapien
19	584	12.9	1203	1 MGR5_RAT	P31424 rattus norv
20	580.5	12.8	1212	1 MGR5_HUMAN	P41594 homo sapien
21	574.5	12.7	1199	1 MGR1_RAT	P23385 rattus norv
22	572.5	12.7	1194	1 MGR1_HUMAN	Q13255 homo sapien
23	474	10.5	999	1 MGR1_MOUSE	O09630 caenorhabdi
24	214	4.7	961	1 GBR1_HUMAN	Q94835 homo sapien
25	212	4.7	960	1 GBR1_MOUSE	Q94v18 mus musculu
26	198.5	4.4	991	1 GBR1_RAT	O92044 mus musculu
27	175	3.8	1103	1 CYGF_BOVIN	O02740 bos taurus
28	173	3.8	940	1 GBR2_RAT	O88871 rattus norv
29	167.5	3.7	941	1 GBR2_HUMAN	O75899 homo sapien
30	159.5	3.5	1108	1 CYGF_HUMAN	P51841 homo sapien
31	151.5	3.3	1108	1 CYGF_RAT	P51842 rattus norv
32	135.5	3.0	1323	1 NME4_RAT	O65645 rattus norv
33	130	2.9	938	1 NMZ1_HUMAN	O05586 homo sapien

34	129	2.9	1323	1 NME4_MOUSE	Q03391 mus musculu
35	128	2.8	938	1 NMZ1_RAT	P35439 rattus norv
36	125.5	2.8	938	1 NMZ1_MOUSE	P35438 mus musculu
37	118.5	2.6	1099	1 CYA7_MOUSE	P51829 mus musculu
38	117.5	2.6	1206	1 VGLM_RVRY	P03518 rlf1 valley
39	116	2.6	1336	1 NME4_HUMAN	O15399 homo sapien
40	115	2.5	986	1 CYGR_ARBPV	P11528 arabacia pun
41	112.5	2.5	1197	1 VGLM_RVRY	P21401 rlf1 valley
42	109	2.4	1110	1 CYGD_BOVIN	P55203 bos taurus
43	108.5	2.4	814	1 AD15_HUMAN	Q13444 homo sapien
44	108.5	2.4	918	1 GIK1_HUMAN	P39086 homo sapien
45	108	2.4	836	1 GIK1_MOUSE	Q60934 mus musculu

ALIGNMENTS

RESULT 1
ID CASR_HUMAN STANDARD: PRT: 1078 AA.
AC P41180: Q13912; Q16379; Q16108; Q16109; Q16110;
DT 01-FEB-1995 (Rel. 31, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Extracellular calcium-sensing receptor precursor (Casr) (Parathyroid
DE Cell calcium-sensing receptor).
GN CASR OR GPRC2A OR PCAR1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Pearce S.H.S., Thakker R.V.;
RL Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
RN [2]
RP TISSUE-Parathyroid;
RC MEDLINE=95279439; PubMed=7759551;
RA Garrett J.E., Capuano I.V., Hammerland L.G., Hung B.C., Brown E.M.,
RT Hebert S.C., Nemeth E.F., Fuller F.;
RL "Molecular cloning and functional expression of human parathyroid
RL J. Biol. Chem. 270:12919-12925(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RX MEDLINE=95408281; PubMed=7677761;
RA Aida K., Koishi S., Tawata M., Onaya T.;
RT "Molecular cloning of a putative Ca(2+)-sensing receptor cDNA from
RL human kidney.";
RL Biochem. Biophys. Res. Commun. 214:524-529(1995).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=96343808; PubMed=8756555;
RA Fretchel M., Zink-Lorenz A., Hollloschi A., Hafner M., Flockert V.,
RT Raue F.;
RT "Expression of a calcium-sensing receptor in a human medullary
RT thyroid carcinoma cell line and its contribution to calcitonin
RT secretion.";
RL Endocrinology 137:3842-3848(1996).
RN [5]
RP SEQUENCE OF 643-908 FROM N.A.
RX MEDLINE=96193893; PubMed=8613532;
RA Birle D.D., Ratnam A., Mauro T., Harris J., Pillai S.;
RT "Changes in calcium responsiveness and handling during keratinocyte
RT differentiation. Potential role of the calcium receptor.";
RL J. Clin. Invest. 97:1085-1093(1996).
RN [6]
RP VARIANTS FHH GLN-185; LYS-297 AND TRP-795.
RX MEDLINE=94094324; PubMed=7916660;
RA Pollak M.R., Brown E.M., Chou Y.-H.W., Hebert S.C., Marx S.J.,
RA Steinmann B., Levi T., Seidman C.E., Seidman J.G.;

RT "Mutations in the human Ca(2+)-sensing receptor gene cause familial
RT hypocalcemic hypercalcaemia and neonatal severe
RT hyperparathyroidism." ;
RL Cell 75:1297-1303(1993).
RN [7]
RP VARIANT ADH ALA-127.
RX MEDLINE=95179179; PubMed=7874174;
RA Pollak M.R., Brown E.M., Estep H.L., McLaine P.N., Kitor O., Park J.,
RA Hebert S.C., Seldman C.E., Seldman J.G.,
RT "Autosomal dominant hypocalcaemia caused by a Ca(2+)-sensing receptor
RT gene mutation." ;
RL Nat. Genet. 8:303-307(1994).
RN [8]
RP VARIANTS FHH MET-62; CYS-66; MET-138; GLU-143 AND GLN-227.
RX MEDLINE=95243222; PubMed=7776161;
RA Chou Y.-H.W., Pollak M.R., Brandi M.L., Toss G., Arnyist H.,
RA Atkinson A.B., Papopoulos S.E., Marx S., Brown E.M., Seldman J.G.,
RA Seldman C.E. ;
RT "Mutations in the human Ca(2+)-sensing-receptor gene that cause
RT familial hypocalcemic hypercalcaemia." ;
RL Am. J. Hum. Genet. 56:1075-1079(1995).
RN [9]
RP SEQUENCE OF 1-61 FROM N.A., AND VARIANT FHH ALA-39.
RX MEDLINE=95403641; PubMed=7673400;
RA Aida K., Koishi S., Inoue M., Nakazato M., Tawata M., Onaya T. ;
RT "Familial hypocalcemic hypercalcaemia associated with mutation in the
RT human Ca(2+)-sensing receptor gene." ;
RL J. Clin. Endocrinol. Metab. 80:2594-2598(1995).
RN [10]
RP VARIANTS NSHPT LEU-227 AND TYR-598.
RX MEDLINE=96292293; PubMed=8675633;
RA Pearce S.H.S., Trump D., Wooding C., Besser G.M., Chew S.L.,
RA Grant D.B., Heath D.A., Hughes I.A., Paterson C.R., Whyte M.P.,
RA Thakker R.V. ;
RT "Calcium-sensing receptor mutations in familial benign hypercalcaemia
RT and neonatal hyperparathyroidism." ;
RL J. Clin. Invest. 96:2683-2692(1995).
RN [11]
RP VARIANTS ADHP THR-116; HIS-681 AND SER-806, AND VARIANT SER-851.
RX MEDLINE=96311554; PubMed=8733126;
RA Barton J., Whner K.K., Yanovski J.A., Cunningham A.W., Laue L.,
RA Zimmerman D., Cutler G.B. Jr. ;
RT "Mutations in the Ca(2+)-sensing receptor gene cause autosomal
RT dominant and sporadic hypoparathyroidism." ;
RL Hum. Mol. Genet. 5:601-606(1996).
RN [12]
RP VARIANT FHH ARG-174.
RX MEDLINE=97442275; PubMed=9298824;
RA Ward B.K., Stuckey B.G.A., Gutteridge D.H., Laling N.G., Pullan P.T.,
RA Ratajczak T. ;
RT "A novel mutation (L174R) in the Ca2+-sensing receptor gene
RT associated with familial hypocalcemic hypercalcaemia." ;
RL Hum. Mutat. 10:233-235(1997).
RN [13]
RP VARIANT FHH GLU-557.
RX MEDLINE=21603857; PubMed=11762699;
RA Nakayama T., Minato M., Nakagawa M., Soma M., Tobe H., Aoi N.,
RA Kosuge K., Sato M., Ozawa Y., Kamatsuse K., Kokubun S. ;
RT "A novel mutation in Ca2+-sensing receptor gene in familial
RT hypocalcemic hypercalcaemia." ;
RL Endocrine 15:277-282(2001).
CC -1- FUNCTION: SENSE CHANGES IN THE EXTRACELLULAR CONCENTRATION OF
CC CALCIUM IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A
CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
CC MESSENGER SYSTEM.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- ALTERNATIVE PRODUCTS: 2 isoforms; 1 (shown here) and 2; are
CC produced by alternative splicing.
CC -1- TISSUE SPECIFICITY: FOUND IN KIDNEY, BUT NOT IN BRAIN, LUNG,
CC LIVER, HEART, SKELETAL MUSCLE, OR PLACENTA.
CC -1- DISEASE: DEFECTS IN CASR ARE A CAUSE OF FAMILIAL HYPOCALCAEMIC
CC HYPERCALCAEMIA (FHH) AND NEONATAL SEVERE HYPERPARATHYROIDISM
CC (NSHPT) TWO INHERITED CONDITIONS CHARACTERIZED BY ALTERED CALCIUM

CC	HOMEOSTASIS. THE MUTATIONS REDUCE THE ACTIVITY OF THE RECEPTOR.
CC	PTH AFFECTED INDIVIDUALS EXHIBIT MILD OR MODERST HYPERCALCEMIA.
CC	RELATIVE HYPOCALCEMIA, AND INAPPROPRIATELY NORMAL PTH LEVELS. IN
CC	CONTRAST, NSHPT IS A RARE AUTOSOMAL RECESSIVE LIFE-THREATENING
CC	DISORDER CHARACTERIZED BY VERY HIGH SERMUM CALCIUM CONCENTRATIONS,
CC	SKLETAL DEMINERALIZATION, AND PARATHYROID HYPERTLASIA. IN SOME
CC	INSTANCES NSHPT HAS BEEN DEMONSTRATED TO BE THE HOMOZYGOUS FORM OF
CC	PTH.
CC	-I- DISEASE: DEFECTS IN CASR ARE ALSO THE CAUSE OF AUTOSOMAL DOMINANT
CC	HYPOCALCEMIA (ADH) IN WHICH THE RECEPTOR IS ACTIVATED AT SUBNORMAL
CC	CAL(2+) LEVELS.
CC	-I- DISEASE: DEFECTS IN CASR ARE ALSO THE CAUSE OF AUTOSOMAL DOMINANT
CC	HYPOPARATHYROIDISM (ADHP). ADHP IS CHARACTERIZED BY HYPOCALCEMIA
CC	AND HYPERPHOSPHATEMIA DUE TO INDEQUATE SECRETION OF PARATHYROID
CC	HORMONE. SYMPTOMS ARE SEIZURES, TETANY AND CRAMP.
CC	-I- STIMULABILITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC	-----
CC	THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION
CC	BETWEEN THE SWISS INSTITUTE OF BIOINFORMATICS AND THE EMBL OUTATION -
CC	THE EUROPEAN BIOINFORMATICS INSTITUTE. THERE ARE NO RESTRICTIONS ON ITS
CC	USE BY NON-PROFIT INSTITUTIONS AS LONG AS ITS CONTENT IS IN NO WAY
CC	MODIFIED AND THIS STATEMENT IS NOT REMOVED. USAGE BY AND FOR COMMERCIAL
CC	ENTITIES REQUIRES A LICENSE AGREEMENT (SEE http://www.isdb-sib.ch/annouce/
CC	OR SEND AN EMAIL TO license@isdb-sib.ch).
CC	-----
DR	EMBL; X81086; CAA56690.1; -
DR	EMBL; U20759; AAA86503.1; -
DR	EMBL; U20760; AAA86504.1; -
DR	EMBL; D50855; BAA09453.1; -
DR	EMBL; S83176; AAB46873.1; -
DR	EMBL; S79217; AAB35262.2; -
DR	EMBL; S68032; AAB29413.2; ALT_SEQ.
DR	EMBL; S68033; AAB29414.1; -
DR	EMBL; S68036; AAB29415.1; -
DR	EMBL; S81755; AAD14370.1; -
DR	GeneW: HGNC:1514; CASR.
DR	MIM; 601189; -
DR	MIM; 145980; -
DR	MIM; 239200; -
DR	MIM; 601198; -
DR	InterPro: IPR001828; ANF_receptor.
DR	InterPro: IPR000337; GPCR_Mgr.
DR	Pfam; PF00003; 7tm_3; 1.
DR	Pfam; PF01094; ANF_receptor; 1.
DR	PRINTS; PR00248; GPCRMR.
DR	PROSITE; PS00979; G_PROTEIN_RECPE_F3.1; 1.
DR	PROSITE; PS00980; G_PROTEIN_RECPE_F3.2; 1.
DR	PROSITE; PS00981; G_PROTEIN_RECPE_F3.3; 1.
DR	PROSITE; PS50259; G_PROTEIN_RECPE_F3.4; 1.
KW	G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
KW	Disease mutation; Alternative splicing; Polymorphism.
FT	SIGNAL 1 19
FT	CHAIN 20 1078
FT	DOMAIN 20 612
FT	TRANSMEM 613 635
FT	DOMAIN 636 649
FT	TRANSMEM 650 670
FT	DOMAIN 671 681
FT	DOMAIN 682 700
FT	DOMAIN 701 724
FT	TRANSMEM 725 745
FT	DOMAIN 746 769
FT	TRANSMEM 770 792
FT	DOMAIN 793 805
FT	TRANSMEM 806 828
FT	DOMAIN 829 836
FT	TRANSMEM 837 862
FT	DOMAIN 863 1078
FT	CARBOHD 90 90
FT	CARBOHD 130 130
FT	CARBOHD 261 261
FT	CARBOHD 287 287
FT	CARBOHD 386 386

Query Match 23.6%; Score 1068.5; DB 1; Length 1078;
 Best Local Similarity 30.3%; Pred. No. 8.6e-68;
 Matches 269; Conservative 156; Mismatches 366; Indels 97; Gaps 26;

12 WALL-----HPGTGAPLCSQQLRMKG DYVGLGFLPG---EAEAGLRSRSPSPVCTR 64
 8 WALLALVHTSAYGP---DQRAQKKGDIIIGLGFPIHFVAKAKDDKMS--RPEVEICR 62
 65 FSSNGLLWALAKMAVEELNNKSDLLPGRLGIDYDFPICSEPVYAKMSLFLA--KAGS 122
 63 YNFRFRMLQAMFAIEEINSSPALLPMLTGLYRIFDCNYSKALEXTLSFVAONKIDS 122
 123 RIIAAYCVTQOPRVAVIGPHSSELMVYKGFSEFLMPOVSGASMEILSAETPPS 182
 123 LNDLFCQSEHPISTINAVGATGSGVSTANALIGLYTIPQVSTASSRLSNKQFES 182
 183 FFRVPSDRVOLTAALAEILOEFGNMVVALGSDDEYGRGSLIFSALAAAGCITAHGL 242
 183 FLRTIPNDEHQATAMADIEYFRMNVGCTIAADDYGRPGIEKFEAEERDICIIFSL 242
 243 VPLPADSRGKQVDYHQNQSSVQVYLLFASVHAHAALFNYSISRLSKVAVASDA 302
 243 I---SQYSDDEELQHVVEVIONSTAKYIVFSSGPDDEPLIKETVRNRITGKIWLASEA 298
 303 WLTSDLVAGLPGMAQMGTVLGLQAGQLOHEPQYV-KTHLALATDPAFCALGERE--- 358
 299 WASSSLIAMPOYFNHYGGTIGFALKAGQIPGFRFLKHYHPKKSXVHNGAFKFMETENC 358
 359 -----QGLEDDVVGQRCPCD-----CITLQNVSA-----GLNHQ- 389
 359 HLOEGAKPLPDVDFLRGHEES--GDRFSNSTAPRPLCTGDEINISVETPYIDYTHIRI 416
 390 TFSYAAAYVAQAALHNLQ-----NASGPAQDPVAPMOLLENNYNTLF-HVGL 440
 417 SYNVLAYVSIHALQDITYTCLPGRGLETPNGS-CADIKKVEAMQYLKHLRLHNFNNNGE 475
 441 PLRFSSGNVMEYDLKLM---VMQGSVRLHDVGRFN-GSLRTRL-----KITMHTSD 491
 476 QYTFEBCDLYGNYSIIMHLSPEDEGSTV-FREVGYYNYAKAGGRLLPINEKILMSGFS 534
 492 NOKPVSRCSROQEGQVRR-VKGFHSCYCDVCEAGSYRONDDIACFCQDEMSPER 550
 535 REVFPSCNSRCLAGTRKGIIEGTEPCFCEVECPDGEISDPTDASACKCPDDPMSEN 594
 551 STCRKRSRPLANGEPVALLLLSLALGLVLAALGFPHHRSPLVQASGGLA--- 607
 595 HTSCIAKEIEFLSWTEPFGIALTLFAVAGIFLTAFLVGVIFKFRNTPIYKATNRELSTYL 654
 608 CFGLVCLGLVCL-SVLEPPGQSPARCLAOPLSHLPLTGLCSTFLQAAEIFY--ESEL 664
 655 LPSLLC-----CFSSSLFETIGEPQDTCRLROPAGISIVLCISLIVTNRYLLFEAKI 710
 665 PLS-----WADRLSGCLGPMAMLVLLAMVEVALCTWYLVAPPEVYVTDHMLPTEAL 719
 711 PTFEHRKMMWMLQ-----FLVFLCTFMQIVICVILYLTAPBSRYNQLDEDEIFE 762
 720 VHCRRSVSVGLAHATNATLAFICELTFLVRSGPCGVNARGLTFPMALAFITWVSFV 779
 763 ITCHEGSLMAGELIGYICLLAALICEFFAFKSRKLPENNEKFTTFEMLLFFIWMISFI 822
 780 PLLANVQVLRPAVOMGALLLCVGLILAFLPRCYLIRNPGRLGTPTPE 827
 823 PAYASTYKFSVAEYVIAIILASFLACIFPNKITYIILFNRNTIE 870

RESULT 2

CASR_MOUSE
 ID CASR_MOUSE STANDARD: PRT: 1079 AA.
 AC Q9QY96: 008968: 09QY95: Q9QZ08: Q9R1D6: Q9R1Z2:
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Extracellular calcium-sensing receptor precursor (CASR) (Parathyroid
 DE Cell calcium-sensing receptor).
 GN CASR OR GPRC2A.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_Taxid=10090.
 RN [1]
 RP SEQUENCE FROM N.A., ALTERNATIVE SPLICING, AND TISSUE SPECIFICITY.
 RC STRAIN=C57BL/6; TISSUE=Kidney;
 RX MEDLINE=20092890; PubMed=10625662;
 RA Oda Y., Tu C.-L., Chang W., Crumrine D., Koemueves L., Mauro T.,
 RA Elias P.M., Bikle D.D.;
 RT "The calcium sensing receptor and its alternatively spliced form in
 RT murine epidermal differentiation.";
 RL J. Biol. Chem. 275:1183-1190(2000).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM A), AND TISSUE SPECIFICITY.
 RC STRAIN=Black Swiss X 129/SVJ; TISSUE=Kidney;
 RX MEDLINE=20119279; PubMed=10652312;
 RA Pl M., Garner S.C., Flannery P., Spurney R.F., Quarles L.D.;
 RT "Sensing of extracellular cations in Casr-deficient osteoblasts.
 RT Evidence for a novel cation-sensing mechanism.";
 RL J. Biol. Chem. 275:3256-3263(2000).
 RN [3]
 RP SEQUENCE OF 256-600 FROM N.A. (ISOFORM A), AND TISSUE SPECIFICITY.
 RC TISSUE=Epiphyseal cartilage;
 RX MEDLINE=20043955; PubMed=10579354;
 RA Chang W., Tu C., Chen T.-H., Komuves L., Oda Y., Pratt S.A.,
 RA Miller S., Shoback D.;
 RT "Expression and signal transduction of calcium-sensing receptors in
 RT cartilage and bone.";
 RL Endocrinology 140:5883-5893(1999).
 RN [4]
 RP SEQUENCE OF 507-582 FROM N.A. (ISOFORM A).
 RC STRAIN=NMRI; TISSUE=Brain;
 RA Hildbrandt J., Ammon H.P.T., Wahl M.A.;
 RT Submitted (May-1998) to the EMBL/Genbank/DBJ databases.
 RN [5]
 RP SEQUENCE OF 562-814 FROM N.A.
 RC TISSUE=Kidney;
 RA Moawad T.I., Riccardi D.;
 RT Submitted (May-1999) to the EMBL/Genbank/DBJ databases.
 RN [6]
 RP SEQUENCE OF 646-799 FROM N.A., AND TISSUE SPECIFICITY.
 RX MEDLINE=97231187; PubMed=9076582;
 RA Quarles L.D., Hartle J.E. II, Siddhanti S.R., Guo R., Hanson T.K.;
 RT "A distinct cation-sensing mechanism in MC3T3-E1 osteoblasts
 RT functionally related to the calcium receptor.";
 RL J. Bone Miner. Res. 12:393-402(1997).
 CC -I- FUNCTION: SENSES CHANGES IN THE EXTRACELLULAR CONCENTRATION OF
 CC CALCIUM IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM.
 CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A (SHOWN HERE) AND B; ARE
 CC PRODUCED BY ALTERNATIVE SPLICING.
 CC -I- TISSUE SPECIFICITY: EPIDERMIS, KIDNEY AND CARTILAGE.
 CC -I- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC
 CC EMBL; AF10178; AAD28371.1; -
 CC EMBL; AF10179; AAD28372.1; -
 CC EMBL; AF128842; AAD40638.1; -
 CC EMBL; AF068900; AAC19388.1; -
 CC EMBL; AB027140; BAA77688.1; -

[illegible]

CC CALCIUM IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A
 CC G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 CC MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -2- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: S67307; AAB29171.1; -
 DR PIR: S40476; S40476.
 DR InterPro: IPR001828; ANF_receptor.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00979; G_PROTEIN_RECP_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECP_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECP_F3_3; 1.
 DR PROSITE: PS02529; G_PROTEIN_RECP_F3_4; 1.
 KM G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 1085
 FT DOMAIN 20 613
 FT TRANSSEM 614 636
 FT TRANSSEM 637 650
 FT TRANSSEM 651 671
 FT TRANSSEM 672 682
 FT TRANSSEM 683 701
 FT TRANSSEM 702 725
 FT TRANSSEM 726 746
 FT TRANSSEM 747 770
 FT TRANSSEM 771 793
 FT TRANSSEM 794 806
 FT TRANSSEM 807 829
 FT TRANSSEM 830 837
 FT TRANSSEM 838 863
 FT TRANSSEM 864 1085
 FT DOMAIN 91 91
 FT CARBOHYD 131 131
 FT CARBOHYD 262 262
 FT CARBOHYD 288 288
 FT CARBOHYD 401 401
 FT CARBOHYD 447 447
 FT CARBOHYD 469 469
 FT CARBOHYD 489 489
 FT CARBOHYD 542 542
 FT CARBOHYD 595 595
 FT CARBOHYD 1085 1085
 SQ SEQUENCE 1085 AA; 121170 MW; 5066D8C9CD13E47 CR664;
 Query Match 23.4%; Score 1060; DB 1; Length 1085;
 Best Local Similarity 30.0%; Pred. No. 3.5e-67;
 Matches 269; Conservative 155; Mismatches 362; Indels 112; Gaps 26;

DB 182 KSLRTIPNDHQATAMADIEFRMMWGTIAADDDYGRGIGKFPREAREIDICIDS 241
 QY 241 GLVPLPRADSRICKVDVYHQVNOSSVOVYLLPASHAAHALNYSISSLRVWVYAS 300
 DB 242 ELI-----SOYSEDEKIOQVVEIONSTAKYIVVSSGDEPLIKEIYRRTITGRIMLAS 297
 QY 301 EAMLTSLYMLKPMQAMQGVYGLQCAQIHEEPQVY-KTHLALADDPACSL----- 354
 DB 298 EAMASSLIAMPEYFHHVGTIGGLAQIPGFRFELQVHPKRSYHNGAKFEWETFE 357
 QY 355 -----GEREGL-----EEDVGOQPCQDCITLON 380
 DB 358 NCHIQEGAKGGLPVDTLRGHEEGAGRLNSPTAFRLCTGEENISSVEPYMDYTHLR- 416
 QY 381 VSAGLNHHQTFSSVYAAYVAQALHNTLQC-----NASCPADQPKFWOLNNYNN 432
 DB 417 -----ISYNYLAVYSIAHALQDITYCIPRGLEFNGS-CADIKKVEAMOVYKILRH 467
 QY 433 LTFHYG-GLPLRFPSGVDMEYDLKLM---VMQSVPLHDVGRFN-GSLRTERL----- 483
 DB 468 LNFYSNMGEOYTFDECGDLAAGNYSITNMHLSPEDESIY-FKEVGYYNYAKKGERLFI 526
 QY 484 -KIRWHTSDNCKPVSRCSQCEQVRR-VKGFHSCCYDCVCEAGSYRONPDDIACITFC 541
 DB 527 EKILMSGFSREVPSPNSRDCIAGTRKGIIEGPCECEYCECPGEYSDETDASACDC 586
 QY 542 GQDEWSPRSRRCRRRRRFLANGEPVLLILLSLALGLVLAALGLFYHHRSPLVQA 601
 DB 587 PDDFWSMNMHTSCIAKELFELSWTEPFIATLRAVLGIFLTAFLVGLIFRMTPIYKA 646
 QY 602 SGGPIA---CEGLVGLVCL-SVLLFPQSPANCAQOPLSHPLTGCTSTLEQAAE 657
 DB 647 TNRELSTILLESLC-----CRESSLFIPEGQDTCRLRQAFGISFLCISCLIVKINR 702
 QY 658 IFV--ESELPLS-----WADRLSGCLRGPMAMLVLLAMLVEVALCTWYIAVAPPEVTD 710
 DB 703 VLVAFEAIFPSFRHKWGLNIQ-----FLVFLCTFMQIYALIMLTAPPSYRN 754
 QY 711 WHMLPTEAL-VHCRTSRVSGLAHATNATLAFICLFTFLVRSQPGCYNARAGITFAML 769
 DB 755 -HELEDELIFITCBGSLMALGFLIGYCLIAACFFPAFRSKRIPEFENAKRTTESML 813
 QY 770 AYEFTWVSFVLNVOQVFLPVAVOMGALLICVGLIAFHLPRCYLLMRPGLTPE 827
 DB 814 IFFLWISFIRAVASTYKRFVSAVEVLAIIAASGILACIFPNKYIILFPRNTIE 871
 RESULT 4
 CASR_RAT
 ID CASR_RAT STANDARD; PRT; 1079 AA.
 AC P48442;
 DT 01-FEB-1996 (Rel. 33, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Extracellular calcium-sensing receptor precursor (CaSR) (Parathyroid
 DE Cell calcium-sensing receptor).
 OS CASR OR GPRC2A OR PCAR1.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC SPRAIN-Sprague-Dawley; TISSUE=Kidney outer medulla;
 RA MEDLINE=95116508; PubMed=7816802;
 RA Riccardi D., Park J., Lee W., Gamba G., Brown E.M., Hebert S.C.;
 RT "Cloning and functional expression of a rat kidney extracellular
 RT calcium/polyvalent cation-sensing receptor.";
 RL Proc. Natl. Acad. Sci. U.S.A. 92:131-135(1995).
 RN [2]
 RP SEQUENCE OF 1-294 FROM N.A.
 RC STRAIN=Wistar;
 RX MEDLINE=95241465; PubMed=7724534;

RA Ruat M., Snowman A.M., Snyder S.H.;
 RT "Calcium sensing receptor: molecular cloning in rat and localization
 to nerve terminals."
 RL Proc. Natl. Acad. Sci. U.S.A. 92:3161-3165(1995).
 CC -1- FUNCTION: SENSE CHANGES IN THE EXTRACELLULAR CONCENTRATION OF
 CALCIUM IONS. THE ACTIVITY OF THIS RECEPTOR IS MEDIATED BY A
 G-PROTEIN THAT ACTIVATES A PHOSPHATIDYLINOSITOL-CALCIUM SECOND
 MESSENGER SYSTEM.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation
 at the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC
 DR EMBL: U01034; AAC52149.1; -
 DR EMBL: U02089; AAC52195.1; -
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMRG.
 DR PROSITE: PS00979; G_PROTEIN_RECP_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECP_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECP_F3_3; 1.
 DR PROSITE: PS50259; G_PROTEIN_RECP_F3_4; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal.
 FT SIGNAL 1 19
 FT CHAIN 20 1079
 FT DOMAIN 20 612
 FT TRANSMEM 613 635
 FT DOMAIN 636 649
 FT TRANSMEM 650 670
 FT DOMAIN 671 681
 FT TRANSMEM 682 700
 FT DOMAIN 701 724
 FT TRANSMEM 725 745
 FT DOMAIN 746 769
 FT TRANSMEM 770 792
 FT DOMAIN 793 805
 FT TRANSMEM 806 828
 FT DOMAIN 829 836
 FT TRANSMEM 837 862
 FT DOMAIN 863 1079
 FT CARBOHYD 90 90
 FT CARBOHYD 130 130
 FT CARBOHYD 261 261
 FT CARBOHYD 287 287
 FT CARBOHYD 386 386
 FT CARBOHYD 446 446
 FT CARBOHYD 468 468
 FT CARBOHYD 541 541
 FT CARBOHYD 594 594
 SQ SEQUENCE 1079 AA; 120867 MW; D7664550361E9736 CRC64;

Query Match 23.3%; Score 1054; DB 1; Length 1079;
 Best Local Similarity 30.3%; Pred. No. 9.1e-67;
 Matches 269; Conservative 163; Mismatches 363; Indels 92; Gaps 28;

9 ISLWALL-HFGTGAPLCLISQOLRMKGDYVGLFPLG---EAEAGLRSTRPSSPYCTR 64
 8 LALLALAMHSSAYGP---DRAOKKGDITLGGFPIHFGVAADQDOKS--REPSVEICR 62
 65 FSSNGLLMALAMKAAVEINNKSDLPGLRIAGYDLFPTCSEPVYAMSLMFLA--KAGS 122
 63 YNFRGFRWLQAMIFAIKEINSSPLLPNMNLTGYRIPTCTWVSVALENTATISFYAQNKIDS 122
 123 RDIAACNTQYQYPRVLAVYGPHSSELAMVTGKFFSEFLMPQVSYGASMSLLSARETFPS 182

DB 123 LNFDEFNCSEHIPSTIAVVGAGSGSVAVANLGLFIYDQVSYASSSRLSKNKYKS 182
 QY 183 FFRVPSDRVQLAAALLOEPGNWVAALGSDDEYRQGLSTIFSALAAARGICIAEGL 242
 DB 183 FLRTIPDEHQATRAMADITFRNNWGTIAADDIDPGIEFRKFEAREEDICIDSE 242
 QY 243 VPLPRADSRIGKVQDVLHGVNOSVQVLLFASVBAHALFNYSISSRLSPKVVAVSEA 302
 DB 243 I-----SQYSDDEEIQVVEVYQNSAKVIYVSSGPDLEPIKEIYARNITGRMLASEA 298
 QY 303 WLISDLVWGLPGMAQMGTVIGFLQRCQALHEFPQYV-KTHLALATDPAFCSALGER--- 357
 DB 299 WASSSLTAMPEYFHVVGITGFGLKAQIGFREFLOKVPKRSVHNGFAKEFWEETPNC 358
 QY 358 --EGLGE-----EDVVGORQCQD-----CITLQNYSA-----GLNHD-TF 391
 DB 359 HLDGAKGPIRPVDTFVASHHEGGRNLNLSSTARPLCTGDEINSVTPMDYHEHLRIST 418
 QY 392 SVYAAVYSVAQALNLTQC-----NASGCPADPKPMQLNNYNTLF-HVGGIPL 442
 DB 419 NVYLAVYSIAHALQDIYTCLPGRGLFTNGS-CADIKVEAMOVYLKHLRHLNFTNMGEQV 477
 QY 443 RFDSSGNVDMEDYDLKLM---VMQSVPRLDVGRFN-GSLRTERL-----KIRNHTSDNQ 493
 DB 478 TFDECGDLVGNYSIINMHLSPEDGSIV-FKEVGYNYAKRGERLFTNEKILMSGFSRE 536
 QY 494 KPYRSCSGROCEGOVR-VKGFHSCCYDCVCEAGSYRONDDIACFFCODEMSPRST 552
 DB 537 VPSNCRDQAGIRKRIITGEFPCCEYCBPGRGEISGETDASACKCDDEKSNENHT 596
 QY 553 RCFRRSRFLAMEGPAVLLLLLSIALGLVLAALGLFHHRSPLQVQASGGLA---CF 609
 DB 597 SIAKEIEFLAMTEPFGIALTFEVLGIFLTAFLVAGVIFKFRMPYKATNRELSYLLF 656
 QY 610 GYVCLGLVCL-SVLLFPGQSPARCAGQQLSHPLPLGCLSTLEQLAAELFY--ESLPL 666
 DB 657 SLIC---CESSSLEFFGEPODWTCLRQAPAGISFVLCISCLIVKRNVLVFEAKIPT 712
 QY 667 S-----WADRLSGCLRGPMAMLVLLAMLEVALCTWYIVAFPEVYTDWMLPTMAL-V 720
 DB 713 SFHRKMGMLNQ-----FLVFLCTPMOILICITWLTAPSSSRN-HELEDEITFI 763
 QY 721 HCRTRSVSEGLAATNATLAFICELGTFLVRSOPGVCYNBARGLTFEMLAYFITWSEVP 780
 DB 764 TCHGSGSLAGSLIGYCLLAALICFFPAFKSRKLPENENAKFTFEMLIFFIWIISFIP 823
 QY 781 LNAVQYVLAIPYMGALLLCYGIILAFHLPXYLLMRQPGINTPE 827
 DB 824 AVASTYGEVSAVEVIAITLAASFGLLACIFPNKVTITLFRPSRNTIE 870

RESULT 5
 ID MGR2_RAT STANDARD; PRT; 872 AA.
 AC P31421;
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Metabotropic glutamate receptor 2 precursor (MG1UR2).
 GN GEM2 OR GPRC1B OR MGUR2.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_Taxid:10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA MEDLINE=92110002; PubMed=1309649;
 RX Tanabe Y., Masu M., Ishii T., Shigemoto R., Nakanishi S.;
 RT "A family of metabotropic glutamate receptors."
 RL Neuron 8:169-179(1992).
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS

MEDICATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE ACTIVITY.

MAY MEDIANE SUPPRESSION OF NEUROTRANSMISSION OR MAY BE INVOLVED IN SYNAPTOGENESIS OR SYNAPTIC STABILIZATION.

-**I**- SUBCELLULAR LOCATION: Integral membrane protein.
TISSUE SPECIFICITY: IS WIDELY DISTRIBUTED IN THE CNS AND PROMINENT
C EXPRESSION IS SEEN IN GOLDI CELLS OF THE CEREBELLUM AND SOME PARTICULAR NEURONAL CELLS IN OTHER BRAIN REGIONS.
C -**S** SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
C SRONESST, TO MGLUJ3.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outpost at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement ([see http://www.isdb.ch/announce/](http://www.isdb.ch/announce/)) or send an email to license@isdb.slb.ch.

CC EMBL; M92075; ; NOT_ANNOTATED_CDS.
DR PIR; JH0561; JH0561.
DR InterPro; IPR001828; ANF_receptor.
DR InterPro; IPR000337; GPCR_Mgt.
DR Pfam; PF00003; 7tm_3; 1.
DR PRINTS; PR00248; GPCRMGR.
DR PROSITE; PS00979; G.PROTEIN_RECIP_F3_1; 1.
DR PROSITE; PS00980; G.PROTEIN_RECIP_F3_2; 1.
DR PROSITE; PS00981; G.PROTEIN_RECIP_F3_3; 1.
DR PROSITE; PS0259; G.PROTEIN_RECIP_F3_4; 1.
KW G-protein coupled receptor; Transmembrane; Glycoprotein; signal;
MW Multigene family.
KM SIGNAL 1 18 POTENTIAL.
FT CHAIN 1 18 METABOTROPIC GLUTAMATE RECEPTOR 2.
FT DOMAIN 19 872 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 568 590 I (POTENTIAL).
FT DOMAIN 591 604 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 605 625 II (POTENTIAL).
FT DOMAIN 626 636 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 637 655 III (POTENTIAL).
FT DOMAIN 656 679 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 680 700 IV (POTENTIAL).
FT DOMAIN 701 725 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 726 747 V (POTENTIAL).
FT DOMAIN 748 760 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 761 783 VI (POTENTIAL).
FT DOMAIN 784 793 EXTRACELLULAR (POTENTIAL).
FT TRANSSEM 794 819 VII (POTENTIAL).
FT DOMAIN 820 872 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 203 203 N-LINKED (GLCNAc . .) (POTENTIAL).
FT CARBOHYD 286 286 N-LINKED (GLCNAc . .) (POTENTIAL).
FT CARBOHYD 338 338 N-LINKED (GLCNAc . .) (POTENTIAL).
FT CARBOHYD 402 402 N-LINKED (GLCNAc . .) (POTENTIAL).
FT CARBOHYD 547 547 N-LINKED (GLCNAc . .) (POTENTIAL).
SQ SEQUENCE 872 AA; 95773 MW; 1E74CBAD6A4BED9 CRC64;

Query Match 14.8%; Score 670.5; DB 1; Length 872;
Best Local Similarity 27.8%; Pred. No. 8.9e-40;
Matches 256; Conservative 133; Mismatches 385; Indels 147; Gaps 37.

DY 1 MGPPVLTSLMALHPGAPLCISQOQLRMRKGDVVIGLGFLPV---GEAEAGLRSTR 56
DB 4 LLLGLATALLMGVAEGPA----KKVLITGGDLVTGGLFPVHKGGFAEECGPYNEHR 57
DY 57 PSSPVCPTFFSSNGTLMLAKMAVEINNKSLETLGLRGYDFPTCSPEVPAAKRPSLMF 116
DB 58 -----GIQRLEMFLADRINDPHILLGVRLGAIIIDSCSDOTALBALDF 105
DY 117 LAKA-----GSRD--AAVCNTYOYPRIAVTVIGHSESELAMYTGKEFFEELMFOVSX 167
DB 106 VRASLDGRDGSRNICPDOSYATHSDAFIAVYGIVSGYSIVANLRLRFQIQSY 165
DY 168 GASMSLLSAKETFEFSFRIVPSDRVQLTRAAAILDEFGNNWYAIALGSDDETGRGLSIIFS 227

Db	166	ASTAKLSDKSRDYDYPARTVYDPDFQAKAMAEILRFNNMTYVSVASGEGYGETGIAFE	225
Qy	228	ALAAARGCINHEGLVPLPRADBSRLKQVDYVHQVQSSVQVYLLEASVHAHALFNYS	287
Db	226	LEAARNRICVATSEKVRGRANSMRAFAEEVYRLI--OKPSARVAVLTREDEAEILL--A	280
Qy	288	ISSRLSPK-VWVASEAMLTSLVNLGPGMOMGVLEFLRGALDHEEP--QYKTHIAL	344
Db	281	ATQRLMASFTWVASDGMGALESVYVAGSERAAEGIT-----IELASYPLISDFRSTYOSL	334
Qy	345	-----ANDPAFCASALGEREOGLEEDVVGORCPQDCITLQNVASGLNHQTFVSYAAVS	399
Db	335	DPMNNSRNPWF-----REFEERFHCFSPFRQDCAHSLRAVPEQESKIME-VVNAVYA	387
Qy	400	VAQALHN---TLQCNASG-CPADPPYKFWQLLEN-MYNLTFHNGGLP-----LRFSSG	448
Db	388	MAHLAHNNHRAICNTHLDCAMAPVNGRRLYKCFVLNVKFDAPFADTDDEVRFPFRFG	447
Qy	449	NVDMEDYDLKVMQGSVP-RLHDVGRFNGLSRTELRKIRNHT--SDNQKPYSRCSROCOEG	506
Db	448	DGIGRYNIETFLYLRAGSGHYRKQKGYNAEGILTDTSTIPMASPACGLPSRCSCEPLQN	507
Qy	507	QVRRYKGHSCCYDVCVDCAEGSYKQNPDDIACFEGODEMSPENSTRCFRRSRFLAMGE	566
Db	508	EKVSQVPEVEVCWMLCIPQPEYRHL--DEFCADCGGIGYPMNASLTCGCFELPOEYIRMGD	565
Qy	567	PAVLLILLILSLAIGLYALAGLVVHHDSPLVQASGPIACFGLVCLGYCL---SVLL	623
Db	566	AWAAGPVYIACLGALATLFTLVGEVVRHNAPEPVYKASREL-CY--ILLGFFLCYCKTFV	622
Qy	624	FPQGPSPARCIAQGPLSHLPLVTGLSTLFLQA---AEIF-----VESLPLSWADRLSG	674
Db	623	FIAPSTAVCTLRRIGIGTAPSVCYSALEITNRIARIFGAREGAPRPRTISASQVAI	683
Qy	675	CIRGPMALVYLLAMLEVEVALCTMYLVAFP-----PEVYTDHMLPTREALVYCRT	724
Db	683	C-----LALISQDLIVA--ANLVYEAPEGTGETAPERREYVT-----LACNH	723
Qy	725	RSWVSFGIAHATNATLAFELCLGTFVLVRSOPGCYNRARGLTFAMLAYFITVWSVPLIAN	784
Db	724	RDASMTLG-SLAVYNVLLIALCTLYAFKTRKCEPENEKAEKIFGTYTTCIIMLAFLETF--	780
Qy	785	VQVYLRAVQWGAALLIC-----VIGLIAHLPRCYLIMPOGLN-----TPEF	828
Db	781	--YITSSDYRQOTTTCMCYSVLSGSVVLGCLFA--PRLHITLFPQKNVVSHPAPTSPF	835
Qy	829	FLGGPGCDAGOCGNDGNTGNGOG	849
Db	836	-----GSAAPRASANLIG-QG	849
RESULT 6			
MGR2_HUMAN			
ID	MGR2_HUMAN	STANDARD;	PRT; 872 AA.
DC	O14415; O9H3M6;		
DT	01-NOV-1997 (Rel. 35, Created)		
DT	01-NOV-1997 (Rel. 35, Last annotation update)		
DT	15-JUN-2002 (Rel. 41, Last annotation update)		
DE	Metabotropic glutamate receptor 2 precursor (mGluR2).		
GN	GRAMOR OR GPRC1B OR MGUR2.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N. A.		
RX	MEDLINE=95346007; PubMed=7620613;		
RA	Fior P. J., Lindauer K., Puttner I., Ruegg D., Lukic S., Knopfel T.,		
RA	Kuhn R.;		
RT	"Molecular cloning, functional expression and pharmacological		
RT	characterization of the human metabotropic glutamate receptor type		
RL	2.,"		
	Eur. J. Neurosci. 7:622-629(1995).		

[2]
 RN SEQUENCE FROM N.A.
 RA Yasuyuki F., Akiko J.;
 RT "Structure and polymorphisms of the human metabotropic glutamate
 receptor type 2 (mglur2) gene: analysis of association with
 schizophrenia.";
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR IS
 CC MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE ACTIVITY.
 CC MAY MEDULATE SUPPRESSION OF NEUROTRANSMISSION OR MAY BE INVOLVED IN
 CC SYNAPTOGENESIS OR SYNAPTIC STABILIZATION.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: WIDELY EXPRESSED IN DIFFERENT REGIONS OF THE
 CC ADULT BRAIN AS WELL AS IN FETAL BRAIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLUR3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 DR EMBL; L35318; AAA7685.1; -;
 DR EMBL; AB045011; BAB19817.1; -;
 DR Genew; HGNC:4594; GRM2.
 DR MIM; 604099; -;
 DR InterPro; IPR001828; ANF_receptor.
 DR InterPro; IPR000337; GPCR_Mgr.
 DR Pfam; PF00003; 7tm_3; 1.
 DR Pfam; PF01094; ANF_receptor; 1.
 DR PRINTS; PR00248; GPCRMR.
 DR PROSITE; PS00979; G_PROTEIN_RECPE_F3_1; 1.
 DR PROSITE; PS00980; G_PROTEIN_RECPE_F3_2; 1.
 DR PROSITE; PS00981; G_PROTEIN_RECPE_F3_3; 1.
 DR PROSITE; PS50259; G_PROTEIN_RECPE_F3_4; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KM MultiGene family.
 FT SIGNAL 1 18
 FT CHAIN 19 872
 FT DOMAIN 19 567
 FT TRANSSEM 568 590
 FT DOMAIN 591 604
 FT TRANSSEM 605 625
 FT DOMAIN 626 636
 FT TRANSSEM 637 655
 FT DOMAIN 656 679
 FT TRANSSEM 680 700
 FT DOMAIN 701 725
 FT TRANSSEM 726 747
 FT DOMAIN 748 760
 FT TRANSSEM 761 783
 FT DOMAIN 784 793
 FT TRANSSEM 794 819
 FT DOMAIN 820 872
 FT CARBOHYD 203 203
 FT CARBOHYD 286 286
 FT CARBOHYD 338 338
 FT CARBOHYD 402 402
 FT CARBOHYD 547 547
 FT CONFLICT 12 12
 FT CONFLICT 210 210
 FT CONFLICT 496 496
 FT CONFLICT 748 748
 FT CONFLICT 776 776
 FT SEQUENCE 872 AA; 95507 MW; 058608C35C701EBD CRC64;
 Query Match 14.6%; Score 658.5; DB 1; Length 872;
 Best Local Similarity 27.4%; Pred. No. 6.3e-39;
 Matches 252; Conservative 131; Mismatches 381; Indels 155; Gaps 37;

QY 9 ISLWALLHPTGAPCLISQOLRMKGIVYGLFPL-----GEAEAGLRSTRPSSPVCTR 64
 Db 11 LPLMGAVAEGR-----KKVLTLEEDLVGLGFPVHQGGAEDCGPVENHR----- 57
 QY 65 FSSKGLMALMKMAVEEINKKSDLLPRLRGIDYDFDPCSEPVYAMKSLFLAKA----- 120
 Db 58 ---GIQRLKMLTALDINDPHLLPGVRLGAHTLSDCSKDTNALAEALDFVRSLSRG 113
 QY 121 --GSRDI---AAVCNYTQYOPRVLAIVPHSSELAATMGKFFSPFIMQVSYGASMETLS 175
 Db 114 ADGSHNICPDGSRATHGAPPAITGVIGGSYSIVQANILRLFOITQIYASISATLS 173
 QY 176 ARETFPSEFPYPSDRVQVLTAAELDFEGNMWVAALGSDDEYGRGSLISFALAAAGI 235
 Db 174 DKSRDYAFARVPDFPFQAKMAEILRFPMNTYVSTEASDEQGTGEARELEARANI 233
 QY 236 CIAHGLVPLTRADDSRLGKQVDVLAHOVNOSSVOYVLLFASVHAHALFNTSISRLSPK 295
 Db 234 CVATSEKGRAMSRRAFGVYRALL--QKPSARAVLFTSEDAARELL--AASORLNAS 288
 QY 296 -VWVASEAM-LTSDLVMLPGMAQMGTVLFLQRCAGLHEFPQYKTHAL----- 344
 Db 289 FTWVASDQMGALLESVYVAGSEAAE-----GATIELASTIPISDPASTYFQSLDPWN 338
 QY 345 -ATDPAFCSALGEREGLEEDVVGQRCP--QCDCT--TLQNVSAGLNHQTFVYAAVYS 399
 Db 339 NSRNPWFREFEYQR-----FRCSFRORDCAHSLRAVPEQESKIME-VYNAVYA 387
 QY 400 VAQALAHN---TLQCNASG--CPADDPVKNPMLLEN-MNVLTFPHVGLP-----LRPDSG 448
 Db 388 MAHALHNHRLALCPHTTILCLDMRPVNGRLKDKLVNWKDAPRPADTINEVRFDRG 447
 QY 449 NVDMEDYDLKLVWQGVSP-RLHDVGRFNGSLRTERIKLRMHTSDNQKPY--SRCSROCE 505
 Db 448 DGIGYNIETYLKAGSGRYRQKGYMAEGTLDTSLIPW-ASPAGPLAASRCEPCIQ 506
 QY 506 GQVRVRFKPHSCDYCVCEGASGRQNDIDIACTFCGDEMSPERSTCFRRSRFLAWG 565
 Db 507 NEVKSVOGEVCCWCWICPCQPEYRL--DEFTCADCGGYPNMSLTGCFELPOEYIRWG 564
 QY 566 EPAVLLLLLLSLALGLVLAALGLFVHRDSPVQVAGSGPLACFLVGLVCL---SVL 622
 Db 565 DAMAVGPTIACGLALTLFVLGAVVRNATPVYKASREL-CY--ILLGVFILCYCTWF 621
 QY 623 LFPQSPSPARLAQDPLSLPLTGLSTLFLQA--AEIF-----VESELPLSWADRLS 673
 Db 622 IFIAPSPAVCTLRRLGTAFSVCYSALLKTRNIATIFGARGAQRPRFISPAQVA 681
 QY 674 GCLGPMALVYVLLAMVEVALCTIYVAF-----PEVYTDHMLPTALVHCR 723
 Db 682 ICL-----ALISQGLITVA--WLVEAPGTGKETAPERREYVT-----LRCN 722
 QY 724 TRSWSFGLAATNATLAFCLFGLFVRSQPGVCNRRAGLTFPMLAVFTWVSFVPLA 783
 Db 723 HRDMSMLD-SLAYNYVALCTLTAFNTKCPENNENKATIGFTYTTCTLTMLALPLF- 780
 QY 784 NVQVLRPAVOMGALLC-----VLGIIAFAHLPRCLYLKROGLN-----TPEFF 829
 Db 781 ---VYTSDDYVQVTTTMCVSVLSGSVYLGCLFA---PKLIIILFQPKNVVSHRAPISR 834
 QY 830 LGGGPGDAGQNDGNTGNG 848
 Db 835 FGSAARASSSLGGSGSQ 853
 RESULT 7
 MGR6_RAT STANDARD; PRT; 871 AA.
 AC P35349;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metabotropic glutamate receptor 6 precursor (mglur6).

GN GRM6 OR GPRC1F OR MGLUR6.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 RX NCBI_TaxID=10116;
 RN (1)
 RP SEQUENCE FROM N.A.
 RC STRAIN=Sprague-Dawley; TISSUE=Retina;
 RA MEDLINE=93280152; PubMed=8389366;
 RA Nakajima Y., Iwakabe H., Akazawa C., Nawa H., Shigemoto R.,
 RT "Molecular characterization of a novel retinal metabotropic glutamate
 RT receptor mGluR6 with a high agonist selectivity for L-2-amino-4-
 RT phosphonobutyrate."
 RL J. Biol. Chem. 268:11868-11873(1993).
 CC -1- IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
 CC ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: RESTRICTEDLY EXPRESSED IN THE INNER NUCLEAR
 CC LAYER OF THE RETINA.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLUR4.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: D13963; BAA03066.1; -
 DR PIR: A46742; A46742.
 DR InterPro: IPR001828; ANF_Receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_Receptor; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00979; G_PROTEIN_RECP_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECP_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECP_F3_3; 1.
 DR PROSITE: PS0259; G_PROTEIN_RECP_F3_4; 1.
 DR G-Protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Multigene family; Vision.
 FT SIGNAL 1 18
 FT CHAIN 1 871
 FT DOMAIN 19 871 METABOTROPIC GLUTAMATE RECEPTOR 6.
 FT TRANSMEM 19 579 EXTRACELLULAR (POTENTIAL).
 FT DOMAIN 580 602 I (POTENTIAL).
 FT TRANSMEM 603 616 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 617 637 II (POTENTIAL).
 FT TRANSMEM 638 648 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 649 667 III (POTENTIAL).
 FT DOMAIN 668 691 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 692 712 IV (POTENTIAL).
 FT DOMAIN 713 742 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 743 764 V (POTENTIAL).
 FT DOMAIN 765 777 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 778 800 VI (POTENTIAL).
 FT DOMAIN 801 813 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 814 839 VII (POTENTIAL).
 FT DOMAIN 840 871 CYTOPLASMIC (POTENTIAL).
 FT CARBOHYD 290 290 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 445 445 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 473 473 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 561 561 N-LINKED (GLCNAC. . .) (POTENTIAL).
 SQ SEQUENCE 871 AA; 95089 MW; 9E70B4D6A13B1867 CRC64;

Query Match 14.0%; Score 631.5; DB 1; Length 871;
 Best Local Similarity 25.6%; Pred. No 5.1e-37;
 Matches 225; Conservative 146; Mismatches 410; Indels 99; Gaps 31;
 Oy 4 PAVIGLSLWALIHPTGAPLCLSQLRMKGDYVIGLGPPLGAEBAEGLRSRTSPSSPVCT 63

DB 5 PVLILMLAMWLSQAGIA---CGAGSVRLAGLITLGGIPVHARGAAG-----RACG 52
 Oy 64 RF-SSNGILLMALAMKMAVEEINKSDLLPGLRYGYDLPDTCSEPVYAKPSPMF----LA 118
 DB 53 ALKKEQGVHRLPAMLYALDRVNADELLPGLRGARLLDPTCSRDYTALEQALSPFOALIR 112
 Oy 119 KAGSDIAACVYTYQYQ-----RYLAVTIPHSSELAAMVTKFFSPFLMQVSYGASM 171
 DB 113 GRGDDEASVRCPGGVPLRSAPPERVAVAVASASVSIMVANYRLFAIPQISASTA 172
 Oy 172 ELLSARETFPSFRTVPSPDRVOLGTAAEELQFGEQNMVVAALGSDDEYRGQSTIFSALA- 230
 DB 173 PELSDSTRDFSRVPPDSTYQAQAMVDYRALGKNVYSTLASSEGNYSGSEVAAVQISR 232
 Oy 231 AARGICIAHGLVPLPRADSRIGVQVDYLVHVNOS-SVQVLLPASHAAHALFNYSIS 289
 DB 233 EAGGVCIAS--IKIPR--EPKPGEFHKVIRRLMETPNARGLIIFANEDDIRRVLEATRQ 288
 Oy 290 SRLSPK-VWVASEAMUTS-DVMGLPGMAQMGTVLGFQRGQQLHEFPQYVTH----- 341
 DB 289 AMLTGHTLVWGSDSWGSKISPLNLEBEA-VGALT-ILPKRASIDGFQDFEYTRSLENN 346
 Oy 342 ----LALATDPAF---GSALGEREGLE---EDVVGQRCPQDCITLQNVASAGLNHO 389
 DB 347 RNIWFAEFMEENFNCKLSSGQSDSTRKCGEERIGD-----SAVEQEK 394
 Oy 390 TFSVTAAYSAVAALHNTLQCNAG---CPADPYKPMQLENNYNLTFH-VGSLPLRF 444
 DB 395 VQFVIDAVYALAHLSHMQALCPHTGICPAMEPTDGRITLHYTRAVFNAGSAGTPVWF 454
 Oy 445 DSSGNVMEYDL-KLWVGQSVPR--LHDVGFNSLIFERIKIWNHSDNCKPVSRGR 501
 DB 455 NENGDAPGRDTPFOYQATNGSASSGQYQAVQWGADELRLDMEVLMWSDGDPHPSPQSL 514
 Oy 502 QCOEGQVRR-VKGFHSCCYDCVDCAGSYRQPDIDACTFCGQDMSPERSTRCEFRSR 560
 DB 515 PCGPGERKKMVGCV-PCMCHEACD--GYRFGVDEFTGACGDMRPNNHNGCAPTPV 571
 Oy 561 FLAKEPRAVLLLLSLALGVLVALGLFVHNRSPVLAQSGGGLACGLVLCGLVLS 620
 DB 572 RLWSSPMAALPLLAVALGIMATTTIMTFMRHNPTVIRASGRELSTVLLGILPILYAI 631
 Oy 621 VLLFGQSPARCLAOPLSHPLNGCSTLFLQA---AEIVSESLPMADRLSGCLR 677
 DB 632 TFLMAEPCAAICAAARLLIGLITLSALTNNRIYRIEQRKRSVTPPPPTIS---- 687
 Oy 678 GPMWLVLV--LAMLVEVALCTWLVAPPEVYTDHMLPT-----EALVACRTSRWVS 729
 DB 688 -PTSQVITFTGLTSQVGVIAM-LGAOPHSHVIDYEQRFVDPQARGVLCDSMSDSL 745
 Oy 730 FGLAATATNTLAFPLFLGTLFLVRSOPGCYNRARGLTFPMLAFITVWVSFVPLANV---- 785
 DB 746 IGCL-GYSLLAMVTCTVYAIKARGVPEFENKPIGFMYTTCIIMLAFVPIFGTAOSA 804
 Oy 786 -QVTLRAVQMGALLCVGLIAAFHLPRCYLLMBOPGIN 824
 DB 805 EKIVYQTTLTVSLSLSASVSIGMLYPKTYVTLFHPDON 844

RESULT 8
 MGR4_HUMAN STANDARD: PRT; 912 AA.
 AC Q14833;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metabotropic glutamate receptor 4 precursor (mGluR4).
 GN GRM4 OR GPRC1D OR MGLUR4.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 CC NCBI_TaxID=9606;

Query	Match	Best local similarity	23.8%	Pred. NO. 2.5e-36	Mismatches 3/4	Indels 204	Gaps 33
Db	4	PAVLGLSIALMLHPTGATGLCL	-----	SOQLRMKGVDYVGLGFP	42		
Db	2	PKRGKLGW	-----	WAPRLCLLLSLVGPWMPSSLGKPKPHNNNSIRIDGDLTGLGFP	56		
Qy	43	LGAEAEAGLRRTPRSSVCTRTSSNGLLMALAKMAVEELINKSDLLPGLRLGTYDLPDT	102				
Db	57	-----VHGRGSGCKPGCEGLKKEKGIHRLKMLFALDRINNDPDLNPTTGARILDT	108				
Qy	103	CSEPVVAMKPSLMFLAKGSRD-IAAYCN-----YTQOPVVLAVIGPHSSELMAMVTKGF	156				
Db	109	CSRDTHALEGSLITYQALIEKDGTVEVRCGSGCPRLTKPEVNVAVIGAGSSSVSIMANTI	168				
Qy	157	FSFELMPQVSYGASMELLSARETEPSEFFRYPPSDRVOLTTAAELLQEGFWMVVAALGSD	216				
Db	169	LRLFLTPQISYASTAPDLSDNSRYDFEGRVYVSDTYQAQAVDILVRLAKMYVSTVASEG	228				
Qy	217	EYRGQGLSTFSLAAR---GICTAHGELVLPRLPADSRIGAKVODVLHQ-VNOSVOVVL	272				
Db	229	SYGEGVYEAFF-IKSRSDGGVCIQAO--VKIPR--EPKAGEFPKIIIRLLETISNARAVI	282				
Qy	273	LEAS-----VHAHALFNYSISSRLSPKVVVASEAMLTGSDL-VMLPGMAOKGVTLGF	324				
Db	283	IFANDDRIRYLEAAR-----ANQGHFEWMSDSGSLIAPVLIHEEVAE-GAVT-I	334				
Qy	325	LQRCALHEFPQYKTH-----LATATPAFCASALGEEQGLEEDVVGORCPQC	373				
Db	335	LPKRHSVAGFPRIYTSRFLDNNRNIMFAEWEENPHCKL---SRHALKKSIVHKKCTNR	391				
Qy	374	DCITLQNVSAGLNHQTFVSYAAYVSAQAQLHNTLQCNASG---CPAADPYKRWOLLEN	429				
Db	392	ERIGD--SAVEOEKQVFIIDAVYAGMHALHAMRDLCPGRVGLCPRYMDPVDGTOLIKY	449				
Qy	430	MYNLTFH-VGGLPLRFDSGQVNDMEYDLKLVWVGSPVRLHDVRFNGSLTEELKLRMH	488				
Db	450	IRNVFSGIAINPVTENFENSGAPGRYDLYOQLRNDSEAYKVITISWDHDLRIETRMNP	509				
Qy	489	TSDNOKPYSRCSROCOEQQVRR-VKGFSCCYDVCDEAGSYRONPDDICTFPGODEWS	547				
Db	510	GSQQGLPSPISLCPQPERKKYTKGM-PCWMHEPCT--GYQVQVDRYTKCTCPYDMRP	566				
Qy	548	PERSTRCEPRRSRLAWGEPAVLLLLLLSLALGLVLAALGLVYHHNDSPLVQASGPPLA	607				
Db	567	TENRGCRPIPLIKLEWSPMAVILPLFLAVYGIAATLEFVITFVRYNDPIVYKASG----	622				
Qy	608	CFGLVCLGLVCLSVLFFPGQSPARCIAQDPLSHPLTG--CLSTFLQAAE-----	657				
Db	623	-----RELSTVLLAGIFLCATYTFLLMAEEDLGTCSS	653				
Qy	658	---TFVESELPLSNADRLSGCLR-----GPAMLVLLAM-LVEVA	694				

Db 654 LRRFLGLGMSISYALLKTKNRIRYIFEOGKRSVAPRPFISPAQSOLATITFSLSQLGL 713
 QY 695 LCTWIVAVPPEVYVDMMHLPTEALVHCRTSRVSGGLAHANATATATFCFTGLVY--- 751
 Db 714 ICWVFEVVD-PSHSYVDEQDQRTLDPRFAR-----GVAKCDISDLISLCLGLYSMLMV 765
 QY 752 -----RSOPGCVNARGLTFFAMLAFTVWSEVFPPL-----ANYQ 786
 Db 766 TCTVVAIKRTGVPEFENAKPIGFTMYTCIWMATIPRIFFGSGSADKLTYQTITLVYS 825
 QY 787 VVLRPAVQMGALLCVLGLTAAFLHPRCYLKRQPLNTPPE 827
 Db 826 VSLASVSLGML-----YMPKYVIIIFHPQONVPK 855
 RESULT 9
 MGR4_RAT
 ID MGR4_RAT STANDARD: PRT; 912 AA.
 AC P31423;
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metabotropic glutamate receptor 4 precursor (mglur4).
 GN GRM4 OR GPRC1D OR MGLUR4.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Euteria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=92110002; PubMed=1309649;
 RA Tanabe Y., Masu M., Ishii T., Shigemoto R., Nakanishi S.;
 RT "A family of metabotropic glutamate receptors";
 RL Neuron 8:169-179(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=93332699; PubMed=8338667;
 RA O'Hara P.J., Sheppard P.O., Thøgersen H., Venezia D., Haldeman B.A.,
 RA McGlane V., Houamed K.M., Thomson C., Gilbert T.L., Mulvihill E.R.;
 RT "The ligand-binding domain in metabotropic glutamate receptors is
 RT related to bacterial periplasmic binding proteins";
 RL Neuron 11:41-52(1993).
 CC -1 FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLYLATE CYCLASE
 CC ACTIVITY.
 CC -1 SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1 TISSUE SPECIFICITY: IS WIDELY DISTRIBUTED IN THE CNS. PREDOMINANT
 CC EXPRESSION IS SEEN IN THE GRANULE CELLS OF THE CEREBELLUM.
 CC -1 SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLUR6.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: M92077; -; NOT ANNOTATED_CDS.
 DR EMBL: M90518; AAA93190.1; -;
 DR PIR: JH0563; JH0563.
 DR InterPro: IPR001828; ANF_receptor.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPRCMGR.
 DR PROSITE: PS00979; G_PROTEIN_RECP_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECP_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECP_F3_3; 1.
 DR PROSITE: PS00259; G_PROTEIN_RECP_F3_4; 1.

KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 Multi-gene family.
 FT SIGNAL 1 32
 FT CHAIN 33 912
 FT DOMAIN 33 587
 FT TRANSMEM 588 610
 FT DOMAIN 611 624
 FT TRANSMEM 625 645
 FT DOMAIN 646 656
 FT TRANSMEM 657 675
 FT DOMAIN 676 699
 FT TRANSMEM 700 720
 FT DOMAIN 721 750
 FT TRANSMEM 751 772
 FT DOMAIN 773 785
 FT TRANSMEM 786 808
 FT DOMAIN 809 821
 FT TRANSMEM 822 847
 FT DOMAIN 848 912
 FT CARBOHYD 98 98
 FT CARBOHYD 301 301
 FT CARBOHYD 454 454
 FT CARBOHYD 484 484
 FT CARBOHYD 569 569
 FT CONFLICT 124 124
 SQ SEQUENCE 912 AA; 101818 MW; 336430EP19BAB577 CRC64;
 Query Match 13.7%; Score 619.5; DB 1; Length 912;
 Best Local Similarity 24.6%; Pred. No. 3; 8e-36;
 Matches 222; Conservative 145; Mismatches 397; Indels 137; Gaps 31;
 QY 4 PAVLGISLALLHPRG-----GAPICLSOOLRMKGIVYGLFPL-GEAEAGLRRTTP 57
 Db 15 PCLLLSLIYAPVWSSVIGKPKGPHNNS--IRIDGDTTGGLEPVHGRSE----- 63
 QY 58 SSPVCTRF--SSNGILMALMKAVEEINKNSDILPGLRLGYDLFDCSEPVYAMKPSLMF 116
 Db 64 -GRACELEKKEKGIIHLEMLFALDRINDPDLPIITIGARLIDCSNDTALBESLUF 122
 QY 117 LAKAGSRD-IAAYCN-----YQYQPRVLAIVGPHSSSELAAYTGKFFSFLMPQVSYGAS 170
 Db 123 VALLLEKDETEVKGSGSPPIITKPERVVGIVGASGVSIWVANILRIKFIPIQISYAST 182
 QY 171 MELLASRETFPSFRTVPDPDQVLTAAELDFEGNWWAALGSDDEYRGOGSIPISALA 230
 Db 183 APDLSDNSRITDFSRVPSDQVQAQMDIVRLAKNNVYSTLASSESYSEGEVAFIQNS 242
 QY 231 AAR-GICIAHEGLVPLPRADDSRLGKQVDVLAH-VNOSVQVVLFPAS-----VHAH 281
 Db 243 RENGGVCAIQNS--VKIPR--EPRTGEFDKIRLLETSNARGIILIFANEDDIRVLEAAR 298
 QY 282 ALFNYSISRLSPKVVAVSEAM-LTSDLVMLGPMQMTVLGFLORGAOLHEFPQVYKT 340
 Db 299 R-----ANQGHFFWMSGDSWGSKAPVLRLEVAE--GAVT-ILPKRMSVRFDFESS 350
 QY 341 H-----LATLTPAFCSALGERBOGLEEDVVGRCPCQDCDITITLQNSAGLNHQ 389
 Db 351 RTLDNRRNIWPAEFWEDNFHCL--SRHALKSGSHIKCTNRRITGD--SAVQBEK 405
 QY 390 TFSYVAAYVSAVALNHTLQCNASG-----CPAADPKPMQLENNMTLNFH-VGGILPFR 444
 Db 406 VQFVDVAVYAMGALHAMRDLCPRGVGLCPRPDPDQVQLKYINNVFSGIANGPVTF 465
 QY 445 DSSGNVDMEDYDLKLVWQGSVPRLDHVGFRNGSLRTERLKIKNHTSDNOKPYVSCRGROQ 504
 Db 466 NENGDAPEGRIYIOYOLRNGSAEYKIVGSWTDLHLIRIRRMQPGSGOOLPFRISCLPQ 525
 QY 505 EGQVRR-VKGFHSCCYDCDCEAGSYRONPDCLATFCODEMSPPRSRRCFRRRFLA 563
 Db 526 PGRKRTVGM-ACWNCERT--GYOYVDRYTCTCPYDMKPTNRRSCODPIYVKE 582
 QY 564 WGPFAVLILLILLISLALGIVLALGLFVHHRSPLVQASGGPLACGLVGLVCLSVLL 623

Db 583 WDSWAVLPFLAVVGIATLFFVVFVRYNDPIKASGRELSTVLLAGIETCYATFL 642
 QY 624 FPGQSPARCLAQPLSHLPGLSTFLQAEIF-----VESELPWADRLSG 674
 Db 643 MIAEPDGTGSLRIEFLGKMSISYALLKTNRIYIFLEGKRSVAPRISASQLA- 701
 QY 675 CLRGPMALVLLAMLEVALCTWYLYVAPPEVVDNMLPTEALVHCRTSVWSFGIAH 734
 Db 702 -----IFELISLQLLGICVWFVD-PSHSVVDFOQDTLDPFRAR-----GVLK 745
 QY 735 ATNATLAFCEIGETFLV-----RSQPCYNRABLTFRAMLAFTTWSPFPL 782
 Db 746 CDISDLSLCLGLGSMILMTCTVYAIKTRGVPTFENAKRIGFTWTTTCVLAFLPIF 805
 QY 783 -----ANQVYLRAVOMGALLCYGLIAFLPRCYLLMRQGLINTP 826
 Db 806 FGTSQADKLYIQTTTLTVSVLSASVSLGML-----YMPKYIILFHPQDNVP 854
 QY 827 E 827
 Db 855 K 855

RESULT 10
 MGR6_HUMAN STANDARD; PRT; 877 AA.
 ID MGR6_HUMAN
 AC O15303;
 DT 15-JUL-1998 (Rel. 36, Created)
 DT 15-JUL-1998 (Rel. 36, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metabotropic glutamate receptor 6 precursor (mglur6).
 GN GNM6 OR GPRC1F OR MGLUR6.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 NC NCBI_TaxID:9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE-97358610; PubMed-9215706;
 RA Hashimoto T., Inazawa J., Okamoto N., Tagawa Y., Bessho Y., Honda Y.,
 RA Nakanishi S.;
 RT "The whole nucleotide sequence and chromosomal localization of the
 RT gene for human metabotropic glutamate receptor subtype 6.";
 RL Eur. J. Neurosci. 9:1226-1235(1997).
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 CC IS MODIFIED BY A G-PROTEIN THAT INHIBITS ADENYLYLATE CYCLASE
 CC ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STRONGEST, TO MGLUR4.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL: U82083; AAB82068.1; -
 CC Genew: HGNC:4598; GRM6.
 DR DR
 DR MIM: 604096; -
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm.3.1.
 DR Pfam: PF01094; ANF_receptor.1.
 DR PRINTS: PR00248; GPCR_MGR.
 DR PROSITE: PS00979; G_PROTEIN_RECEPTOR_F3_1; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECEPTOR_F3_2; 1.
 DR PROSITE: PS00981; G_PROTEIN_RECEPTOR_F3_3; 1.
 DR PROSITE: PS50259; G_PROTEIN_RECEPTOR_F3_4; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Multigene family; Vision.

FT SIGNAL 1 24
 FT CHAIN 25 877
 FT DOMAIN 25 585
 FT TRANSMEM 586 608
 FT TRANSMEM 609 622
 FT TRANSMEM 623 643
 FT TRANSMEM 644 654
 FT TRANSMEM 655 673
 FT TRANSMEM 674 697
 FT TRANSMEM 698 718
 FT TRANSMEM 719 748
 FT TRANSMEM 749 770
 FT TRANSMEM 771 783
 FT TRANSMEM 784 806
 FT TRANSMEM 807 819
 FT TRANSMEM 820 845
 FT TRANSMEM 846 877
 FT CARBOHYD 296 296
 FT CARBOHYD 451 451
 FT CARBOHYD 479 479
 FT CARBOHYD 567 567
 SQ SEQUENCE 877 AA; 95436 MW; AECDF9LE8DA5994F CRC64;
 Query Match 13.68; Score 617; DB 1; Length 877;
 Best Local Similarity 24.98; Pred. No. 5,4e-36;
 Matches 212; Conservative 151; Mismatches 398; Indels 92; Gaps 31;

QY 29 LRMKGVYVGLGFLPGLGEAEAGLRSTRPSSPYCFRRSSNGLLMALAKMAVEINRKS 88
 Db 33 VRLAGGLTGLGFLPVHARGAAG-----RACGPL---KKEGVNHEALMALDRVNDPE 84
 QY 89 LLPGLRLGYDLFDTCSEPVYAMKPSLMFL-----AKGSRDIAIYC---NTQYQY----R 137
 Db 85 LLPGVRRGARLRDTCSDYDTALAEQALSFVQALIRGRDGDGVRCGGVPLRPAPPR 144
 QY 138 VLAIVGPHSSLEAMVTGKFSFELMPQVSYGASMEILSARETFSPRYPSDQVQLTA 197
 Db 145 VVAVVGASASVSSTIMVANLRLFAIPQISTASTAPELSDSTRYDFSRVVPDPSYQAM 204
 QY 198 AELLQERGMWVVAALGSDDEYRGQSLIFSALA--AANGICIAHGLVPLPRADSRIGKV 256
 Db 205 VDIYRALGMWVSTLASBGNNGEGVEAFVQISAEAGVCAQS--IKIR--EPKGEF 260
 QY 257 QDVLHGVNQS--SVQVLLFASVHAALFNYSSISRLSPK--VWVASAM--LTSQVLWGLP 313
 Db 261 SKVIRLMEIPNARGIILFIANEDIRRVLEAARQANLTGHLMVSDSGWCAKTPISLSE 320
 QY 314 GMAQMGVVLGFLQAGDALHERPYVKH-----LALATPACSKLGEBOGLE 363
 Db 321 DVA-VGAIL-ILPRASIDGDDYETFRSLNNRNIMFPEWENENCKL--TSSGTQS 376
 QY 364 DVVGQRPOCDCTLQVNSAGLNHQTFS-----YAAVYSVAQALHMTLQCNASG-- 414
 Db 377 DDIRKRC-----TGERIGRDSYEBQKQVFIADYALTAHALSHQALCGHRT 427
 QY 415 --CPADQVPRWQLENNYNTLTFH-VGGLPRFDSGNVDMEDYL-KLWVQGSVPR--L 468
 Db 428 GLCPAMEPTGRMLLYITRAVRFNGSAGTFVMEFNEDADAGRYDIFQYQATNSASSGY 487
 QY 469 HDVGRFNGSLRTERLKRIMHTSDQKPVSRCSRCQCGOVRR--VKGHSCCYQDVDEAG 527
 Db 488 QAVGQMETLRDLDEALQWGGDPHEVPSLCSLPCGGERKMKVKGV--PCWCBEACD-- 544
 QY 528 SYRNPDIACITPCGQDEWSPERTSRCFRRSRFLANGEPAVILLILLSLATGLVLAAL 587
 Db 545 GYRQVDEFTCEACPGDMRPPIPNHTGCRPRPVYVLSMSSPMAAPPLLAIVLGIVATTIV 604
 QY 588 GLFYVHNDSPLYQASGGLACFGLVGLVCLVSLVLPFGQSPARCLAQPLSHLPGLTC 647
 Db 605 ATFVRYNNTPIVRASGRELSTVLLTGIFLYAITFLVAWBAVCAARRLFGLGTTLS 664
 QY 648 LSTFLDA---AEIFVSELPWADRLSGCLRPAMLYVL--LAMLVEVALCTWYLYVA 702

Db 665 YSALLTKNRIRIREFGCKRSVTPPEPIS-----PTSQVITTSLSIQVGMIAW-LGA 718
 QY 703 PPEVETVMHMLPT-----EALVHCRTSMWSFGLAHNTNLFELCLGTFIVRSQSG 756
 Db 719 RPHSHVIDEEOQRTVDEQARGLVKDMSDLISGL-CYSLLIHWCTGVAIKARGVE 777
 QY 757 CYNARGLTFAMLAFTVWSFVPLIANV-----OVLPAPVOMGALLICVGIILAAFL 811
 Db 778 TENEKPIGFMTYTCIIMLAFVPIFFGIAQSAEKIYIQTTLTVYSLISASVIGMLYV 837
 QY 812 PRCYLLMRPGIN 824
 Db 838 PRTYVILFHEQN 850
 RESULT 11
 MGR7_HUMAN
 ID MGR7_HUMAN STANDARD: PRT; 915 AA.
 AC Q14831;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metabotropic glutamate receptor 7 precursor (mGluR7).
 GN GRM7 OR GPRC1G OR MGLUR7.
 OS Homo sapiens (Human).
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 CC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=96437220; PubMed=8840028;
 RA Makoff A., Pilling C., Harrington K., Emson P.;
 RT "Human metabotropic glutamate receptor type 7: molecular cloning and
 RT mRNA distribution in the CNS.";
 RL Brain Res. Mol. Brain Res. 40:165-170(1996).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=96141892; PubMed=9473604;
 RA Wu S., Wright R.A., Rokey P.R., Burgett S.G., Arnold J.S.,
 RA Rostock P.R., Jr., Johnson B.G., Schoepf D.D., Belagaje R.M.;
 RT "Group III human metabotropic glutamate receptors 4, 7 and 8:
 RT pharmacological properties in RGT cells.";
 RL Brain Res. Mol. Brain Res. 53:88-97(1998).
 RN [3]
 RP VARIANT PHE-433.
 RX MEDLINE=21095249; PubMed=11163549;
 RA Bolonna A.A., Kerwin R.W., Munro J., Arranz M.J., Makoff A.J.;
 RT "Polymorphisms in the genes for mGluR types 7 and 8: association
 RT studies with schizophrenia.";
 RL Schizophren. Res. 47:99-103(2001).
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
 CC ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN MANY AREAS OF THE BRAIN,
 CC ESPECIALLY IN THE CEREBRAL CORTEX, HIPPOCAMPUS, AND CEREBELLUM.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STROMESEST, TO MGLUR4.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL: X94552; CAA64245.1; -
 CC EMBL: U92458; AAB51763.1; -
 CC Genew; HGNC:4599; GRM7.
 DR MIM; 604101; -
 DR

DR InterPro: IPR001828; ANF receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF000003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS; PR00248; GPCRMR.
 DR PROSITE; PS00979; G_PROTEIN_RECP_F3_1; 1.
 DR PROSITE; PS00980; G_PROTEIN_RECP_F3_2; 1.
 DR PROSITE; PS00981; G_PROTEIN_RECP_F3_3; 1.
 DR PROSITE; PS0259; G_PROTEIN_RECP_F3_4; 1.
 DR G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KM Multigene family; Olfaction; Polymorphism.
 KW SIGNAL.
 FT CHAIN 1 32
 FT DOMAIN 33 915
 FT DOMAIN 33 590
 FT TRANSMEM 591 615
 FT DOMAIN 616 627
 FT TRANSMEM 628 648
 FT DOMAIN 649 654
 FT TRANSMEM 655 675
 FT DOMAIN 676 702
 FT TRANSMEM 703 723
 FT DOMAIN 724 753
 FT TRANSMEM 754 775
 FT DOMAIN 776 788
 FT TRANSMEM 789 810
 FT DOMAIN 811 825
 FT TRANSMEM 826 850
 FT DOMAIN 851 915
 FT CARBOHYD 98 98
 FT CARBOHYD 458 458
 FT CARBOHYD 486 486
 FT CARBOHYD 572 572
 FT VARIANT 433 433
 SQ SEQUENCE 915 AA; 102250 MW; CFF94E06FF74919 CMC64;
 Query Match 13.6%; Score 615.5; DB 1; Length 915;
 Best Local Similarity 24.2%; Pred. No. 7.3e-36;
 Matches 215; Conservative 167; Mismatches 385; Indels 121; Gaps 32;
 QY 6 VGLSLMALHGTGAPLCLSQLRKQDYVIGGLFPLGLAEAGLRSTRSSPYCTGF 65
 Db 20 VLEVLICALAAARQGEWAPHSIRIGDVTGLGGLFPVAKGSGV-----PGDIR 71
 QY 66 SSGGLMALAMKAVEINIKNSDLLPLGLRGLDPTGCEPVAMKPSIMFLAKAGSRDI 125
 Db 72 RENGHRLRMLYALDQINSDPMLPNTGLRILDTCSRDTYALAEOSLTFVQALIQKDT 131
 QY 126 A-AYCNYTOYOP-----RLAVIGPHSEELAMVYGFSEFLMPQVSGAMELISAR 177
 Db 132 SDVRC--TNGEPFVYFKPKRYGVIGAGSSVSIMVANTLRFLQIDISYASTAPRLSD 189
 QY 178 ETPPSEFTVPSDRVOLTAABELDQEGNNWVAALGSDDEYGRQGISFSAALAAARGIC 236
 Db 190 RRYDFFSRVPPDSFQAQAMVDIVKALGNVYSTLASSESGYSGKVESTQISKEGGGLC 249
 QY 237 IAHGELVPLPRDDSLGAVOVLIHQ-VNOSQOVVLELAS-VHAHALFNSSISRLSP 294
 Db 250 IASQVRIPOERKD--RTIDFDRIKQLDTPNSRAVITRANDEDIKOITLAARKRAQVGH 307
 QY 295 KYVAASEAMLT-----SDLVWG-LIPGMAQGVYIGF-----LQGAQLHEFPY 337
 Db 308 FLWVGSNMGSKINPLHGEDIAEGAITIQPKRATVEGDVFTSRTLENRRNVFAFY 367
 QY 338 VATHALADPAFCSLGREGLEDVVGRCPCDCITLQNVASGLHHHQTFS---YV 394
 Db 368 WEENFNCK-----LITSGSKEDTDKRCGO-----ERIGDSNYEDEGKVOFYI 412
 QY 395 AAVYSVAQALHN--TLQGNASG-CPAODPVPMQLENNMYLTFH-VGGLPLRFPSSGN 449
 Db 413 DAVYMAHALHMKNDCLADYRGVCPPEMOAGSKILKIRIVNFGSAGTVPYMKND 472
 QY 450 VDMEDYL-KLAWVQGSVPLHDVGRNGSLRTERLIRWHTSDNCKPVYRCSROCEGOV 508

Best Local Similarity 23.4%; Pred. No. 1,3e-35;
Matches 210; Conservative 159; Mismatches 380; Indels 147; Gaps 29;

QY 26 SQQLRMKGDYVGLGFLPGLGEAEAGLRSTRPSSPVCTRSSNGLMALAMKMAVEEINN 85
D 36 SVSVPFLPDIIIGLFLPVEHKE-----GAPCGPKVYNRGVQRLMALMAYLADRYNN 86
QY 86 KSDLLPGLRLGYDLFDTCSEPVVAMKPSLML-AKAGSRDIAVCNTYQYQPR----- 137
D 87 DPNILPGITIGVHILDTSRPTVALNOSIQFVRASLINMIDTSQVCEADGSSFOLEKNASS 146
QY 138 --VLAVIPHSSELMVWGKFFSPFLMPOVSGASMEILSAREPSPFRFPYPSDRYLT 195
D 147 GPVEGVIGSGSVSLQVLANILRLPHIPOVSPASTAKLSDKTRDLARVPPDPTFSV 206
QY 196 AAELLOFGNMVVAALIGSDDEYGRQGISFSAALAAAGICIAHGLVPLPRAD--DSR 252
D 207 ALVDILKFNMSYVSTISESGYGEIHLKRETERNVCIAVAKEYPSAADKVPDSI 266
QY 253 LGKQVDVIAHQVNOSSVOVLLFASVHAHALEFNISISRLS-PKRVVASEAM-LSLDIYM 310
D 267 ISKLD-----KKPVAKGVVLETRAEADRRLIQAARKNISQPFHMIASDGGKQOKLLE 320
QY 311 GLPGMAQGTVLGFLORGAQLHEFPQY-----VKTHLALA 345
D 321 GLEDIAE-GALTVELQ--SEIIADFDRYMMQLPETNQNPFAEYMEDTFNCLVLSVK 378
QY 346 TDPAFCSALGEREOGLIEDVVGQRCPCDCITLONVSGNLNHQ--TFSYAANYSAQ 402
D 379 PDTNSANSTNKIGVK-----AKTECDYSRLSEKVGEDSKTQFVADYAYAY 430
QY 403 ALHN--TLQCN-----ASGCPADQVVKPMOLLEN-MY 431
D 431 ALHNHNDRCMTQSSQTTETRKHLQSESVWTRKISTDTRKSOACPMANYSDEKEFYNNL 490
QY 432 NLTF-HVGLPLREDSSGNVMEYDLKLVMQ--GSVPLHDVGRNPSLTERLKIWMH 488
D 491 NVSFLDLAGSEVKFDRODGLARVDILNVQROENSSGVYKVGKWMFNGLOLNSTYVWN 550
QY 489 TSDNOKPVSRCROCEQVAVKGFHSCCYDCVDCENAGSTRONDIACTFCGDEWSP 548
D 551 -KETEQPTSAQSLPCEVMIRKQOG-DTCWCWICSCSEFEVY--DEFTCKDCGGLMPY 606
QY 549 ERSTRCFRRSRFLAMGSPAVLILLLLSLALGLVLAALGLFVHHRDPIVOASGPIA- 607
D 607 ADKLSCTALDIOYKMNLSFLILPMAIIFGIALTSIVIVFAKNHDPVLRASGRELSTY 666
QY 608 --CFGLVCLGLVC-LSVLLFPGOPSPARCLAQO-----PLSHPLTGLCISTLEFLQ 654
D 667 TILFGI----LVYCYNTPALLAKPTIGSCVLRGIGVGFISALLTKNRISRIHS 722
QY 655 AAELFVESELPLSWADRLSGLRGFWAMLVYLLAMLVVALCTWYLVAPPEVYTDWMML 714
D 723 ASKS-----AQRK--YISPOSOVYITSLDIAOVITLMTMVAVEEPTGFYYPD 770
QY 715 PTEALVHCRTSRWSVFGALHATNATLAFCLGFLVSPGOCYRARGLEFAMAYFLT 774
D 771 RREVLTKRKID-KSFLFSQLXMLLITICITIAKTKRIENNESFIFGTMTTTCII 829
QY 775 WVSFVPLLANYVLRPAVOMGALLCYLGILAA-----FHLPRCYLLMRPGIN 824
D 830 WLAEPVIFYFT-----GNSYEVQTTTLCISISLSASVALVCLSPRVYLLVHPDKN 881

RESULT 13
MGR3_HUMAN
ID MGR3_HUMAN STANDARD; PRT; 877 AA.
AC Q14832;
DE 01-NOV-1997 (Rel. 35, Last sequence update)
DE 01-NOV-1997 (Rel. 35, Last sequence update)
DE 15-JUN-2002 (Rel. 41, Last annotation update)
DE Metabotropic glutamate receptor 3 precursor (mGluR3).
GN GRM3 OR GPRC1C OR MGLUR3.

OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxId:9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Brain;
RX MEDLINE-96437205; PubMed-8840013;
RA Makoff A., Volpe F., Leichuk R., Harrington K., Emson P.;
RT "Molecular characterization and localization of human metabotropic
RT glutamate receptor type 3";
RL Brain Res. Mol. Brain Res. 40:55-63(1996).
CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
CC ACTIVITY.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
CC STRONGEST, TO MGLUR2.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; X77748; CA54796.1; -;
CC Genew; HGNC:4595; GRM3.
CC MIM; 601115; -;
CC InterPro; IPR001828; ANF_Receptor.
CC InterPro; IPR000337; GPCR_Mgr.
CC Pfam; PF00003; 7tm_3; 1.
CC Pfam; PF01094; ANF_receptor; 1.
CC PRINTS; PR00248; GPCRMR.
CC PROSITE; PS00979; G_PROTEIN_RECP_F3_1; 1.
CC PROSITE; PS00980; G_PROTEIN_RECP_F3_2; 1.
CC PROSITE; PS00981; G_PROTEIN_RECP_F3_3; 1.
CC PROSITE; PS0259; G_PROTEIN_RECP_F3_4; 1.
CC KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
CC MultiGene family.
CC FT SIGNAL 1 20
CC FT CHAIN 21 877
CC FT DOMAIN 21 574
CC FT TRANSMEM 575 597
CC FT DOMAIN 598 611
CC FT TRANSMEM 612 632
CC FT DOMAIN 633 643
CC FT TRANSMEM 644 662
CC FT DOMAIN 663 686
CC FT TRANSMEM 687 707
CC FT DOMAIN 708 732
CC FT TRANSMEM 733 754
CC FT DOMAIN 755 767
CC FT TRANSMEM 768 790
CC FT DOMAIN 791 800
CC FT TRANSMEM 801 826
CC FT DOMAIN 827 807
CC FT CARBOHYD 207 207
CC FT CARBOHYD 290 290
CC FT CARBOHYD 412 412
CC FT CARBOHYD 437 437
CC SEQUENCE 877 AA; 98619 MW; 66F28663CE35F740 CRC64;
Query Match 13.4%; Score 604.5; DB 1; Length 877;
Best Local Similarity 24.9%; Pred. No. 1,3e-35;
Matches 225; Conservative 151; Mismatches 372; Indels 157; Gaps 35;

QY 25 LSQQLRMKGDYVGLGFLPGLGE-----AERAGLRSTRPSSPVCTRSSNGLMALAMKMAV 80
D 27 LRRLEIKIGDLYVLGLGFLPINKGTGTECGRINER-----GIORLEMLPFI 74
QY 81 EFINNKSDLLPGLRLGYDLFDTCSEPVVAMKPSLML-AKAGSRDIAVCNTYQYQPR--- 135

DB 75 DEINDDYLLGVKGLHLLTSCSDTYALQSOSEFVRSALTKYDEAYMCPDGSYAIQ 134
 QY 136 --PRVLA-VIGPHSSELMTGKFFSFILMPQVSGASMELLASRETPSEFRYPSDRV 192
 DB 135 NIPPLINGVIGSSVSIQVIANLLRLQIPIQIISASTSALSKSRIDYFARYTPDPFY 194
 QY 193 QLTAAELLOEFGNMVNAALSDDEYSGROGISALSALAANGICIAHEGLVPLPRADDSR 252
 DB 195 QAKMAEILREFNMVTVSTVASEGDYGETIEAFQEARLNICIATAEKGRSNIKRSY 254
 QY 253 LGKVDVLYHOVNOSSVOYVLLF-----ASHVAHAALFNYSISSRLSPKVVASEA 302
 DB 255 DSVIRELLQKFN---ARVYVLFMSDSRELIILASRNABF-----TWASDG 300
 QY 303 WLTSDVYGLPGMAQMGTVLGLQGAO-LHEFPQYKT-----HLATDPAF----- 350
 DB 301 WGAQESIITKSEBHAAYGAIIT--LELASPVYQFDPYFQSLMPYNNH-----KNPWRDWE 354
 QY 351 ----CSALGEREOGLEEDVYQORCPQDCITLQNVSAQINHQFTS-----YVAAVYVAQ 402
 DB 355 QKFOCSLQNKRNH-----RRV--CD-----KHLAIDSSVYQESKIMFYVNAVYAMAH 400
 QY 403 ALH---NTIQNAGS-CPADDPVKWOLLEN-MYNLTFHVGLP-----LRPSSGNV 450
 DB 401 ALHKMORTLCNNTIKLCDAKILQOKKLYKYLKINTAPPNPKDADSIKEDTFEDG 460
 QY 451 DMEYDLKIMWQGSVYPLRDYGRFNGSLRTERLKIWMHSTNOKVRSRQCOGQYR 510
 DB 461 MGRNVNFQVWVGKYSYLAKYGHNAETLSLVNSIHW--SNNSVYTSQSDPCAPAKENKN 518
 QY 511 VKGHSCCYDVCDEAGSYRQNPDDIACFGQODEMSPERSTRCFRRSRLANGEPAYL 570
 DB 519 MQPBDVCCWICIPCE--PYEYLADEFTCMDCSGQWMPADLTGCTDEDEDYIRMEDAWAI 576
 QY 571 LLLLLLALDLVLAALGLFVHNDSPLYOASGGPLACFGVGLVGLVSLVLPFGQSP 630
 DB 577 GPVITACIGFECTCMVYVTFKHNNTPLVAKSGRELCTYLLFGVGLSCMTFFIAKSP 636
 QY 631 ARCLAQPLSHPLTGLSTLFLQ---AEFVESELPLSWADRLSGCLRG-----P 679
 DB 637 VICALRRILGLSSFAICYSALLTKNCLARIF-----DYKNAQARKFISP 683
 QY 680 WAMLVILLAM-LVEVALCTWLVAFPRPVYVDMHMLPE---ALVHCRTSRMVSGLAHA 735
 DB 684 SSOVFICLGLLVQIVMVSVMILLEAPG--TRRYTLAKRETVILKICVWKO--SSMLISLT 740
 QY 736 TNATLAEFCFGLTFLVRSOPGCYNRARGLTFAMLAFTTWSEFVLANVOVLRPANOM 795
 DB 741 YDVLIVLICVYAAKTRKCPENFEAKFIFTMTTTCIIMLFLPIF---YVSSDYRV 796
 QY 796 GALLIC-----VLGIILAAFLHPRCYLLMRPGMLTPEF-----FLGGGPDAGQ 840
 DB 797 QTTMCISIVSLSGFVVLGLCPFA---PKVHILLPQOKNVVHRLHNFVSIGGTTYSQ 853
 QY 841 MDGNT 845
 DB 854 SSAST 858

RESULT 14
 MGR7_RAT
 ID MGR7_RAT STANDARD; PRT; 915 AA.
 AC P35400;
 DT 01-JUN-1994 (Rel. 29, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Metabotropic glutamate receptor 7 precursor (mglur7).
 GN GMR7 OR GPRCIG OR MGLUR7.
 OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;

RP SEQUENCE FROM N.A.
 RC STRAIN-Sprague-Dawley; TISSUE-Brain;
 RX MEDLINE=94117433; PubMed=828685;
 RA Okamoto N., Hori S., Akazawa C., Hayashi Y., Shigemoto R.,
 RA Mizuno N., Nakanishi S.;
 RT "Molecular characterization of a new metabotropic glutamate receptor
 RT mglur7 coupled to inhibitory cyclic AMP signal transduction.";
 RL J. Biol. Chem. 269:1231-1236(1994).
 RP SEQUENCE FROM N.A.
 RC STRAIN-Sprague-Dawley; TISSUE-Olfactory bulb;
 RX MEDLINE=94195260; PubMed=8145723;
 RA Saugstad J.A., Kinzie J.M., Mulvihill E.R., Segerson T.P.,
 RA Westbrook G.L.;
 RT "Cloning and expression of a new member of the L-2-amino-4-
 RT phosphonobutyric acid-sensitive class of metabotropic glutamate
 RT receptors.";
 RL Mol. Pharmacol. 45:367-372(1994).
 CC -1- FUNCTION: RECEPTOR FOR GLUTAMATE. THE ACTIVITY OF THIS RECEPTOR
 CC IS MEDIATED BY A G-PROTEIN THAT INHIBITS ADENYLATE CYCLASE
 CC ACTIVITY.
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -1- TISSUE SPECIFICITY: WIDELY DISTRIBUTED THROUGHOUT THE BRAIN.
 CC -1- SIMILARITY: BELONGS TO FAMILY 3 OF G-PROTEIN COUPLED RECEPTORS.
 CC STROMEST, TO MGLUR4.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: D16817; BA004092.1; -;
 DR EMBL: U06832; AA20653.1; -;
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR PRINTS: PR00248; GPCRMR.
 DR PROSITE: PS00979; G_PROTEIN_RECPT_1.
 DR PROSITE: PS00980; G_PROTEIN_RECPT_F3_1.
 DR PROSITE: PS00981; G_PROTEIN_RECPT_F3_3; 1.
 DR PROSITE: PS50259; G_PROTEIN_RECPT_F3_4; 1.
 KW G-protein coupled receptor; Transmembrane; Glycoprotein; Signal;
 KW Multigene family; Olfaction.
 FT SIGNAL 1 32
 FT CHAIN 1 915
 FT DOMAIN 33 590
 FT TRANSMEM 591 615
 FT DOMAIN 616 627
 FT TRANSMEM 628 648
 FT DOMAIN 649 654
 FT TRANSMEM 655 675
 FT DOMAIN 676 702
 FT TRANSMEM 703 723
 FT DOMAIN 724 753
 FT TRANSMEM 754 775
 FT DOMAIN 776 788
 FT TRANSMEM 789 810
 FT DOMAIN 811 825
 FT TRANSMEM 826 850
 FT DOMAIN 851 915
 FT CARBOHYD 98 98
 FT CARBOHYD 458 458
 FT CARBOHYD 486 486
 FT CARBOHYD 572 572
 SQ SEQUENCE 915 AA; 102231 MW; F28AFc4C645AaC2 CRC64;
 Query Match 13.4%; Score 604.5; DB 1; Length 915;
 Best Local Similarity 23.9%; Pred. No. 4,4e-35;

Db 77 DEINKDNVLLPGVKGHILDTCSRDYALQSLFEVRAASLTKYDEAEYMCPOGSAIOE 136
QY 136 --PRYLA-VIGPHSELAMVWGKFFSFILMPOVSGASMEILLASARETFPSFRTYPSDRV 192
Db 137 NIPLLIAGVIGSYSSVSIQVIANLRLFPQISTASTAKISDKSRDYAFARTVPPDFY 196
QY 193 QLTAAELLQEFGNMVAALGSDDEYRGQGLISFALAANGICIAHGLVPLPRADDSR 252
Db 197 OAKAAEILRFENNTYVSTVASEGDEYGETEAFQEARLNINICITATAEKVGRSNIKRSY 256
QY 253 LGRVODVLAHQVQSSVOVYLLFASVHAHALEFNYSISRLSPK-VWVASEAM-LTSDLV 310
Db 257 DSVIRELLQKRN--ARYVVLFRSDSDREL--AAANRVNASFTWVASDGVAGQESIVK 311
QY 311 GLPGNAQGVTLGFLQGAOLHEFPQYKT-----HLALATDPAF-----CSALG 355
Db 312 GSEHVAIYCAITLELASH--PYRQFDRYQSLNPNYNN--RNFWRDPEWQKFOCSLQ 365
QY 356 EREGLEDVYVGORCPQCCITLQNVASAGLNHQTFSYAAVYSVAQALH--NTLQCN 412
Db 366 KRNH-----ROYCDKHLAIDSSNYE--QESKIMFVNAVYAMAHALHKMQRTLCPT 415
QY 413 SG-CPAODPVKFWOLL-ENMTNLTFHVGLP-----LRDSSGNVMEYDLKIMWOG 463
Db 416 TKLCDAMKILDGKLYKEYLLKINFTAPFNPKGADSIYKFDTEGDMGRYVFNMQOTG 475
QY 464 SVPLRHVDYRFGNSLRTERLKI RMTSDNOKPVSRQSGOEGOYRVRKGFHSCCYD 523
Db 476 GKYSYKRGHNAETISLDVDSIH--SRNSVPTSCSDPCAPENKKNMQPGDYCCWITCIP 533
QY 524 CEASYSRONPDDIACFCGODEWSPERSTRCFRRSRFLANGEPVALLLLSLALGV 583
Db 534 CE--PYEYLVDEFQCMDCGPGMPTADLSGCYNLPEDYIKMEDAWAIGPVTIACGFLCT 591
QY 584 LAALGFVHHHDSPLYQASGGPLACFGVCLGYCLVSLVLPFGOPSPARCLAQPLSHLP 643
Db 592 CIVITVFYKHNTPLVYKASGRELICYLLFGVSLSYCMTFPFIKPSVICALRRLGLGTS 651
QY 644 LTGCLSTFLQA--AEIFVESELPLSWADRLSGCLRG-----PMAMLVILLAM-LV 691
Db 652 FAICYSALETKTNCIARIF-----DGVKNGAQRPKFTSPSSQVFIICGLIIV 698
QY 692 EVALCTWYLVAPPEVVTDMHMLPTE--ALVHCRTSMVSFGLAHATNATLAFICFLGT 748
Db 699 QIVWVSWLLILETPG--TRRYTLEPKRETVILKCNVKD--SSMLISLTYDVVLVIICTYVA 755
QY 749 FLVRSQPCYRARGLTFAMLAFTWVSFPLANOVVLRAVQMGALLC----- 801
Db 756 FKTRKCPENENKAFIGFTMTTCTIIMAFLPIF--YVTSDDYRVQVTTWCISVSLSG 811
QY 802 --VLGIILAAFHLPRCYILMROPGLN 824
Db 812 FVVLGCLFA--PKVHIYVLPQPKN 833

Search completed: May 19, 2003, 09:50:32
Job time: 16.0911 secs

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:48:15 ; Search time 38.8189 Seconds
(without alignments)
4522.333 Million cell updates/sec

Title: US-09-927-315-15
Perfect score: 4524
Sequence: 1 MGPAYLGLSLMALHPGTG.....GPGDAQQNDNGTNGQKHE 852

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 671580 seqs, 206047115 residues

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_21:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phage:*
- 10: sp_plant:*
- 11: sp_rotent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriap:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3255	71.9	858	11	0923K1	0923K1 rattus norv
2	3229	71.4	858	11	091VA4	091VA4 mus musculu
3	3227.5	71.3	858	11	0925A4	0925A4 mus musculu
4	3226.5	71.3	858	11	0925D9	0925D9 mus musculu
5	3221.5	71.2	858	11	0925D8	0925D8 mus musculu
6	3221.5	71.2	858	11	0923K0	0923K0 mus musculu
7	1208.5	26.7	840	11	0920R8	0920R8 rattus norv
8	1199.5	26.5	842	11	0920R5	0920R5 mus musculu
9	1195.5	26.4	842	11	0925I5	0925I5 mus musculu
10	1188.5	26.3	842	11	0923J9	0923J9 mus musculu
11	1180.5	26.1	842	11	0929P6	0929P6 mus musculu
12	1077.5	23.8	763	4	08TDJ9	08TDJ9 homo sapien
13	1060	23.4	940	13	090WU6	090WU6 sparus aura
14	1049	23.2	940	13	073635	073635 fugu rubrip
15	1014.5	22.4	839	4	08TE23	08TE23 homo sapien
16	1009.5	22.3	877	13	09PW88	09PW88 carassius a

17	1002.5	22.2	843	11	0920R7	0920R7 rattus norv
18	1000.5	22.1	843	11	0925I4	0925I4 mus musculu
19	999.5	22.1	843	11	0923I8	0923I8 mus musculu
20	933.5	20.6	880	13	073639	073639 fugu rubrip
21	925	20.4	864	13	073637	073637 fugu rubrip
22	876.5	19.4	912	11	0704I0	0704I0 mus musculu
23	845	18.7	844	13	093552	093552 carassius a
24	835.5	18.5	848	13	093553	093553 carassius a
25	828	18.3	868	13	073636	073636 fugu rubrip
26	827.5	18.3	856	13	073638	073638 fugu rubrip
27	760	16.8	875	13	073640	073640 fugu rubrip
28	616.5	13.6	855	11	0704O9	0704O9 mus musculu
29	612.5	13.5	1218	13	0902P3	0902P3 oncorhynch
30	605.5	13.4	983	11	0629I6	0629I6 rattus norv
31	601.5	13.3	877	4	08TBH8	08TBH8 homo sapien
32	601.5	13.3	879	11	090YS2	090YS2 mus musculu
33	592.5	13.1	779	11	035269	035269 rattus norv
34	583.5	12.9	1156	13	098UC6	098UC6 gallus gall
35	583.5	12.9	1188	13	098UC5	098UC5 gallus gall
36	583.5	12.9	1242	13	098UC4	098UC4 gallus gall
37	575.5	12.7	1199	11	09EPV6	09EPV6 mus musculu
38	555.5	12.3	866	11	035268	035268 rattus norv
39	550.5	12.2	977	13	09PWE1	09PWE1 ictalurus p
40	516.5	11.4	852	11	035192	035192 mus musculu
41	512	11.3	850	11	035189	035189 mus musculu
42	507.5	11.2	457	4	08TDJ8	08TDJ8 homo sapien
43	503	11.1	667	11	035267	035267 rattus norv
44	503	11.1	768	11	035266	035266 rattus norv
45	498.5	11.0	803	11	035191	035191 mus musculu

ALIGNMENTS

RESULT 1

ID	Query Match	Score	DB	Length	Result
0923K1	71.9%	3255	DB 11	858	0923K1
AC	0923K1				PRELIMINARY; PRT; 858 AA.
DT	01-DEC-2001				(TREMBLrel. 19, Created)
DT	01-DEC-2001				(TREMBLrel. 19, Last sequence update)
DT	01-JUN-2002				(TREMBLrel. 21, Last annotation update)
DE	Sweet taste receptor TIR3.				
GN	TASIR3.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				SEQUENCE FROM N.A.
RP	STRAIN-WISTAR.				
RA	Neelson G., Hoon M.A., Chandrasekar J., Zhang Y., Ryba N.J.P.,				
RA	Zuker C.S.,				
RL	"Mammalian Sweet Taste Receptors.";				
RL	Cell 0:0-0(2001).				
RN	[2]				SEQUENCE FROM N.A.
RP	STRAIN-WISTAR.				
RC	STRAIN-WISTAR; TISSUE=CIRCUVALLATE PAPILLA;				
RA	MEDLINE=21927605; Pubmed=11917125;				
RA	Li X., Staszewski L., Xu H., Durick K., Zoller M., Adler E.,				
RT	"Human Receptors for Sweet and Umami Taste.";				
RT	Proc. Natl. Acad. Sci. U.S.A. 99:4692-4696(2002).				
DR	EMBL; AY032620; AAK51601.1; -				
DR	EMBL; AF456324; AAK10636.1; -				
DR	InterPro; IPR001828; ANF_receptor.				
DR	InterPro; IPR000337; GPCR_Mgr.				
DR	Pfam; PF00003; 7tm_3; 1.				
DR	Pfam; PF01094; ANF_receptor; 1.				
DR	PROSITE; PS00980; G_PROTEIN_RECPT_F3_2; UNKNOWN_1.				
DR	PROSITE; PS50259; G_PROTEIN_RECPT_F3_4; 1.				
KW	Receptor.				
SO	SEQUENCE. 858 AA; 94845 MW; 98890DAF75973B80 CRC64;				

QY	Db	Accession	Species	Length	Score	E-value	Bit Score	Identities	Positives	Gaps	Conserved Domains	Annotations
305	Db	TSDIWLTPNIAKGVGVGFQJRGALLPERSHVEYTHLALAAPACASL-NAELDLEH	Human	363	11.1	1e-11	11.1	100%	100%	0%	Putative taste receptor	
365	QY	VVGGRPCPCDCITLQNSAGL-----NHQTSYAAVYVVAQALNHTLQCNASGCP	Human	416	10.8	1e-10	10.8	100%	100%	0%	Putative taste receptor	
364	Db	VMGQRCRCDDIMQLNSSLGQLQMSAGQHLQIFATYAAVSVQAQALNHTLQCNASHQ	Human	423	10.7	1e-10	10.7	100%	100%	0%	Putative taste receptor	
417	QY	AADPYKPMQLLENNYNLTFHVAGLPLRPDSSGVNDMEYDLKLWQGSVPRLDHVGRENG	Human	476	10.6	1e-10	10.6	100%	100%	0%	Putative taste receptor	
424	Db	VSEHPLPMQLLENNYNNMSFHARDLTLLQFDAGEVNDMEYDLKMWQSPPTVLTHTVGTENG	Human	483	10.5	1e-10	10.5	100%	100%	0%	Putative taste receptor	
477	QY	SLRTSLKILRMHTSDNCKPYSRGRCQGEQVRRVGFHSCYDCVDCAGSYRONDDI	Human	536	10.4	1e-10	10.4	100%	100%	0%	Putative taste receptor	
484	Db	TLQLQASMTY--FGQVPPYSQCRQCKDQVRRVGFHSCYDCVDCAGSTRKHDPF	Human	541	10.3	1e-10	10.3	100%	100%	0%	Putative taste receptor	
537	QY	ACTFCGQDQSPERSTRCFRRSRFTLAMEPAVLLLLLLLSTALGLVALGLGVHRRDS	Human	596	10.2	1e-10	10.2	100%	100%	0%	Putative taste receptor	
542	Db	TCTPCNQDQSPERSTRCFRRSRFTLAMEPAVLLLLLLLSTALGLVALGLGVHRRDS	Human	601	10.1	1e-10	10.1	100%	100%	0%	Putative taste receptor	
597	QY	PLNGASGRPLACDELNLGLVCLSVLLFPQPSAPARLQAQPSLHPPLRGCLSTFLQAA	Human	656	10.0	1e-10	10.0	100%	100%	0%	Putative taste receptor	
602	Db	PLVQASGGSQCFELICIGLFCULSVLLFPGRPSSASCLQAQPAHNLPTGCLSTFLQAA	Human	661	9.9	1e-10	9.9	100%	100%	0%	Putative taste receptor	
657	QY	EIFVESELPLSMARLSCGLRGPAWLVVLLAMLEVALCTWYLVAFPEVYVDHMLPT	Human	716	9.8	1e-10	9.8	100%	100%	0%	Putative taste receptor	
662	Db	ETFESELPLSMAMWLCISYLRGLMAMLVVLLAFVEALCAWLLAFPEPVYVDHMLPT	Human	721	9.7	1e-10	9.7	100%	100%	0%	Putative taste receptor	
717	QY	EALVHCRTKRSWVSGLAHATNATLAFLCFTGLTVLVSQPCYRNARGLFAMLATYITWV	Human	776	9.6	1e-10	9.6	100%	100%	0%	Putative taste receptor	
722	Db	EVLHCRTKRSWVSGILVHITAMLAFLCFGLTGLTVLVSQPCYRNARGLFAMLATYITWV	Human	781	9.5	1e-10	9.5	100%	100%	0%	Putative taste receptor	
777	QY	SEVFLANVOYVLPAVOMGLILICVGLIIAFLPLPCYLLMPOGLNPEFLG-----	Human	831	9.4	1e-10	9.4	100%	100%	0%	Putative taste receptor	
782	Db	SEVFLANVOYVPAVOMGLILVCAILGILVTHLPKCYVLLMPLKINTOEFLGNNAK	Human	841	9.3	1e-10	9.3	100%	100%	0%	Putative taste receptor	
832	QY	832-----GGPGDAQOND 842	Human	842	9.2	1e-10	9.2	100%	100%	0%	Putative taste receptor	
842	Db	842 AADENSGGGEAQAQCHNE 858	Human	858	9.1	1e-10	9.1	100%	100%	0%	Putative taste receptor	

Query Match	71.3%;	Score 3227.5;	DB 11;	Length 858;
Best Local Similarity	72.4%;	Pred. No. 2.6e-250;		
Matches 620;	Conservative 67;	Mismatches 156;	Indels 13;	Gaps 4;

QY	5	AVLGSLTMAILLHPGICGAPLCLSLQOLRMGDVYLGGLFLPLGEHEBGLSLRRRSPSPVCTR	64
Db	5	ALMGSLTAAFLTELGGASLCLSLQOLFKAGODVTLGGFLPGSTEEETLQORAPNSTLCLR	64
QY	65	FSSNGLLMAALAKMAVVEEINNKSDLLPGRLGLYDLEFDCSEPVVAMKPSIMFLAAGSRD	12
Db	65	FSPGLGLFAMAKMAVEEINNKSSALLPGRLGLYDLEFDCSEPVVAMKSSIMFLAAVSGQS	12
QY	125	IAAYNTQYOYOPRYLAVIGPHSSELAMVTGKFFSFLMPQVSYGASMELLSARETFPPSEF	18
Db	125	IAAYNTQYOYOPRYLAVIGPHSSELALITGKFFSFLMPQVSYGASMSRLSRETFPPSEF	18
QY	185	RTVPEDRYOLTAALAELEBEFNWVAALGSDDEYXROGHSIFISALAAARGICIAHEGLVP	24
Db	185	KVPSEDRYOLQALAVYLLQLNFSSMWVAALGSDDDYDREGISITSSLANARGICIAHEGLVP	24
QY	245	LPRAADSRIGKYQDYLHOVNOSSVOYVLLFASVAHAHALENTSISSRUSPKVWVASEAWL	30
Db	245	QHDITGGQOOLQKLDVLYROVNOSSKQYVVVYLFASARAVYSLFSTSIHHGLSPKWVASEWL	30
QY	305	TSDDLWMLPGMAOMCEVULGFIQRGALOHFEPOUYVTHLALADDPFCSALGEREGLEED	36
Db	305	TSDDLWMLPNAIKVGIVLGFILQRGALLPBFSHYVEHLLALADDPFCSALNAEIDLEEH	36
QY	365	VVGORPCPODCITLQONVSAGL-----NNHQTESVYAAYVSVAQALHNTLQCNASGCP	41
Db	364	VVGORPCPODCITLQONLSSGILLQNLASQDLHHQIFATYAAYVSVAQALHNTLQCNVSHCH	42
QY	417	AODPYKPMQLEENMTNLFPHVGGELPLRPDSSGNVDMEDYLKATWMOGSYPRLHDYGRFNG	47
Db	424	VSEHVLPMQLEENMTNMSFHARDLTLQFABEGNVMEYDLKMWMOSSPTPVLYHTVGTENG	48
QY	477	SLRTERLKRMTSTNOKRVSRCSROCOGQVRRYKGFHSCCYDCVDCAGSYRONPDDI	53
Db	484	TLQLOQSKMYM--PENQVPRVSGSCQKXQGRYRRYKGFHSCCYDCVDCAGSYRKHDPDF	54
QY	537	ACTFFGCODEMSPERSTKCFRRRSRPLANGEPYVLLLLLSTALGLVTLAALGLFYNNHRS	59
Db	542	TCTPCNODMSPKSTACLPBRPKFLAMGEPRVYLSLLLCITLVGLTALAALGLSVHHNMS	60
QY	597	PLVQASGGFLAFEGVYGLVCLVLSVLFPGQSPSPARCLAQOPRLSHLPLTGCLSTLYLQA	65
Db	602	PLVQASGGSGQCFGLICLGLFCLSVLFLPFRGRSSASCLAQOPMAHPLPLTGCLSTLYLQA	66
QY	657	ELFVSESELPLSMADRLSGCLRGPMALVYLLAMLVEALCTWYLYVAFPEPVVTDMMHMLPT	71
Db	662	ELFVSESELPLSMANNMLCSYLRGIMLMLVYLLATFVEAALCAVYLYAFPEPVVTDMSVLPRT	72
QY	717	EALVYCRTRPSWVSPGLAHNTNATLAFICPLGFIIVRSOGCCVNRARGFLFAMLAFTITVY	77
Db	722	EALVYCHRSWVSLGVLVHTNMLAFELCLGFIIVLOSQGRKYNRARGFLFAMLAFTITVY	78
QY	777	SFVPLILANQVYLRAVQMGALLCVLGLTILAFHPLRCYLIRKQGLNTPPEFLGGPGPD	83
Db	782	SFVPLILANQVAYQPAVQMGAILVICALGITLVFHFPLKCYVLLMLPLRNTQEEFLGRNAKK	84
QY	837	AGQMDNGTNGNGKHE 852	
Db	842	AADENSG--GGEAGE 855	

RESULT	4			
ID	Q925D9	PRELIMINARY;	PRT;	858 AA.
Q925D9				
AC	Q925D9;			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, last sequence update)			
DT	01-JUN-2002 (TREMBLrel. 21, last annotation update)			
DE	Putative sweet taste receptor family 1 member 3.			
GN	TAS1R3.			
OS	Mus musculus (mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]
 RX SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J;
 MEDLINE=21225287; PubMed=11326277;
 RA Max M., Shanker Y.G., Huang L., Rong M., Campagne F., Liu Z.,
 Weinstein H., Damak S., Margolske R.F.;
 RT "Tas1r3, encoding a new candidate taste receptor, is allelic to the
 RT sweet responsiveness locus Sac.";
 RL Nat. Genet. 28:58-63(2001).
 DR EMBL: AF368024; AAK5536.1; -
 DR MGD: MGI:1933547; Tas1r3.
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECIP_F3_2; UNKNOWN_1.
 DR PROSITE: PS50259; G_PROTEIN_RECIP_F3_4; 1.
 KW Receptor.
 SQ SEQUENCE 858 AA; 94633 MW; 4E416824FDB478F8 CRC64;

Query Match 71.3%; Score 3226.5; DB 11; Length 858;

Best Local Similarity 72.4%; Pred. No. 3.2e-250;
 Matches 620; Conservative 66; Mismatches 155; Indels 13; Gaps 4;

OY 5 AVLGSLMALHPGTGAPLCLSQOLRMKGDVYLGLEPLGAEAGLRSTRPSPVCTR 64
 DB 5 AIMGLSLAAFLGMASTLCSQOFKAGDYILGGLPTEATLNORQAPSPICNR 64
 OY 65 FSSNGLLMALAKMAVEINNKSDLLPGLRGYDLPTCSEPVYAMKPSLMFLAKGSRD 124
 DB 65 FSPGLFLAMAKMAVEINNKSGALLPGLRGYDLPTCSEPVYAMKPSLMFLAKGSRD 124
 OY 125 IAAACNTQYQPRVLAIVGHSSSLAMVTGKFFSFILMPQVSYASAMDRISDRTPPSFF 184
 DB 125 IAAACNTQYQPRVLAIVGHSSSLAMVTGKFFSFILMPQVSYASAMDRISDRTPPSFF 184
 OY 185 RTVPSDRVQTLAAELLQEFGMNVAALGSDDEXRGGLSTFSALAAARGICIAHEGLVP 244
 DB 185 RTVPSDRVQTLAAELLQEFGMNVAALGSDDEXRGGLSTFSALAAARGICIAHEGLVP 244
 OY 185 RTVPSDRVQTLAAELLQEFGMNVAALGSDDEXRGGLSTFSALAAARGICIAHEGLVP 244
 DB 185 RTVPSDRVQTLAAELLQEFGMNVAALGSDDEXRGGLSTFSALAAARGICIAHEGLVP 244
 OY 245 LPRADSRGLKQVDYLVHQVNSQVYVLLFASVAHAALFNYSISSRLSPKVVWVASEAML 304
 DB 245 LPRADSRGLKQVDYLVHQVNSQVYVLLFASVAHAALFNYSISSRLSPKVVWVASEAML 304
 OY 245 QHDTSGOGLGVLDYLVHQVNSQVYVLLFASVAHAALFNYSISSRLSPKVVWVASEAML 304
 DB 245 QHDTSGOGLGVLDYLVHQVNSQVYVLLFASVAHAALFNYSISSRLSPKVVWVASEAML 304
 OY 305 TSDLVMLPGMAQNGVYLVGLQGAQLHEPPQYKTHLATATDPAFCASALGEROGLEED 364
 DB 305 TSDLVMLPGMAQNGVYLVGLQGAQLHEPPQYKTHLATATDPAFCASALGEROGLEED 364
 OY 305 TSDLVMLPGMAQNGVYLVGLQGAQLHEPPQYKTHLATATDPAFCASALGEROGLEED 364
 DB 305 TSDLVMLPGMAQNGVYLVGLQGAQLHEPPQYKTHLATATDPAFCASALGEROGLEED 364
 OY 365 VVGQRCPQDCITLQNVASGL-----NHQTFSVYAAYVSYAVALHNTLQCNASGCP 416
 DB 365 VVGQRCPQDCITLQNVASGL-----NHQTFSVYAAYVSYAVALHNTLQCNASGCP 416
 OY 364 VMGQRCPQDCITLQNVASGL-----NHQTFSVYAAYVSYAVALHNTLQCNASGCH 423
 DB 364 VMGQRCPQDCITLQNVASGL-----NHQTFSVYAAYVSYAVALHNTLQCNASGCH 423
 OY 417 AQPVRKWMOLLNNYNTFVHGLPLRPDSSGNVDMEDYDKLVWQGSVPRLHVGREGN 476
 DB 417 AQPVRKWMOLLNNYNTFVHGLPLRPDSSGNVDMEDYDKLVWQGSVPRLHVGREGN 476
 OY 424 VSEHVLFWOLLNNYNTFVHGLPLRPDSSGNVDMEDYDKLVWQGSVPRLHVGREGN 483
 DB 424 VSEHVLFWOLLNNYNTFVHGLPLRPDSSGNVDMEDYDKLVWQGSVPRLHVGREGN 483
 OY 477 SLRTERLKIMHTSDNOKPVSRCSROCEQVRRKGFHSSCCYDCOCEASQONPDI 536
 DB 477 SLRTERLKIMHTSDNOKPVSRCSROCEQVRRKGFHSSCCYDCOCEASQONPDI 536
 OY 484 TLQLOQSKMHW--PGNVPVYSCQSCQCKDQVRYKGFHSSCCYDCOCEASQONPDI 541
 DB 484 TLQLOQSKMHW--PGNVPVYSCQSCQCKDQVRYKGFHSSCCYDCOCEASQONPDI 541
 OY 537 ACFFCGODEMSPSTRCFRRSRFLAMGEPVALLLLSLALGLVLAALGLFVHHNDS 596
 DB 537 ACFFCGODEMSPSTRCFRRSRFLAMGEPVALLLLSLALGLVLAALGLFVHHNDS 596
 OY 542 TCFPCNODQMSPEKSTACLRPRRPFLLAMGEPVALLLLSLALGLVLAALGLFVHHNDS 601
 DB 542 TCFPCNODQMSPEKSTACLRPRRPFLLAMGEPVALLLLSLALGLVLAALGLFVHHNDS 601
 OY 597 PLVQASGGLACGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLV 656
 DB 597 PLVQASGGLACGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLV 656
 OY 602 PLVQASGGLACGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLV 661
 DB 602 PLVQASGGLACGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLVGLV 661
 OY 657 EIFVESGLPLSMADRLSGCLRGPMVAVLLAMVEVALCTWTVYARPEVYDWMHLP 716
 DB 657 EIFVESGLPLSMADRLSGCLRGPMVAVLLAMVEVALCTWTVYARPEVYDWMHLP 716
 OY 662 EIFVESGLPLSMADRLSGCLRGPMVAVLLAMVEVALCTWTVYARPEVYDWMHLP 721
 DB 662 EIFVESGLPLSMADRLSGCLRGPMVAVLLAMVEVALCTWTVYARPEVYDWMHLP 721

OY 717 EALVHCRTSRWVSEGLAHATNATLAFICLGTFLVRSQPCGNRARGITFAMLAFTWV 776
 DB 722 EVLEHCHRSVNSLGVHITAMALAFICLGTFLVRSQPCGNRARGITFAMLAFTWV 781
 OY 777 SFVPLLANVQVYVLPVAVOMGALLICVIGITLAAFLPRCYLLMRPGLMTPPEFYGGGSD 836
 DB 782 SFVPLLANVQVYVLPVAVOMGALLICVIGITLAAFLPRCYLLMRPGLMTPPEFYGGGSD 841
 OY 837 AOGQNDGNTGNGKHE 852
 DB 842 AADENSG--GGEAAOE 855

RESULT 5

ID 0925D8 PRELIMINARY; PRT; 858 AA.
 AC 0925D8;
 DT 01-DEC-2001 (TREMBlrel. 19, Created)
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
 DE Putative sweet taste receptor family 1 member 3.
 GN TAS1R3.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OC NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/SVEV;
 RX MEDLINE=21225287; PubMed=11326277;
 RA Max M., Shanker Y.G., Huang L., Rong M., Campagne F., Liu Z.,
 Weinstein H., Damak S., Margolske R.F.;
 RT "Tas1r3, encoding a new candidate taste receptor, is allelic to the
 RT sweet responsiveness locus Sac.";
 RL Nat. Genet. 28:58-63(2001).
 DR EMBL: AF368025; AAK5537.1; -
 DR MGD: MGI:1933547; Tas1r3.
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PROSITE: PS00980; G_PROTEIN_RECIP_F3_2; UNKNOWN_1.
 DR PROSITE: PS50259; G_PROTEIN_RECIP_F3_4; 1.
 KW Receptor.
 SQ SEQUENCE 858 AA; 94533 MW; 2012102855D7EB60 CRC64;

Query Match 71.2%; Score 3221.5; DB 11; Length 858;

Best Local Similarity 72.3%; Pred. No. 8e-250;
 Matches 619; Conservative 67; Mismatches 157; Indels 13; Gaps 4;

OY 5 AVLGSLMALHPGTGAPLCLSQOLRMKGDVYLGLEPLGAEAGLRSTRPSPVCTR 64
 DB 5 AIMGLSLAAFLGMASTLCSQOFKAGDYILGGLPTEATLNORQAPSPICNR 64
 OY 65 FSSNGLLMALAKMAVEINNKSDLLPGLRGYDLPTCSEPVYAMKPSLMFLAKGSRD 124
 DB 65 FSPGLFLAMAKMAVEINNKSGALLPGLRGYDLPTCSEPVYAMKPSLMFLAKGSRD 124
 OY 125 IAAACNTQYQPRVLAIVGHSSSLAMVTGKFFSFILMPQVSYASAMDRISDRTPPSFF 184
 DB 125 IAAACNTQYQPRVLAIVGHSSSLAMVTGKFFSFILMPQVSYASAMDRISDRTPPSFF 184
 OY 185 RTVPSDRVQTLAAELLQEFGMNVAALGSDDEXRGGLSTFSALAAARGICIAHEGLVP 244
 DB 185 RTVPSDRVQTLAAELLQEFGMNVAALGSDDEXRGGLSTFSALAAARGICIAHEGLVP 244
 OY 185 RTVPSDRVQTLAAELLQEFGMNVAALGSDDEXRGGLSTFSALAAARGICIAHEGLVP 244
 DB 185 RTVPSDRVQTLAAELLQEFGMNVAALGSDDEXRGGLSTFSALAAARGICIAHEGLVP 244
 OY 245 LPRADSRGLKQVDYLVHQVNSQVYVLLFASVAHAALFNYSISSRLSPKVVWVASEAML 304
 DB 245 LPRADSRGLKQVDYLVHQVNSQVYVLLFASVAHAALFNYSISSRLSPKVVWVASEAML 304
 OY 245 QHDTSGOGLGVLDYLVHQVNSQVYVLLFASVAHAALFNYSISSRLSPKVVWVASEAML 304
 DB 245 QHDTSGOGLGVLDYLVHQVNSQVYVLLFASVAHAALFNYSISSRLSPKVVWVASEAML 304
 OY 305 TSDLVMLPGMAQNGVYLVGLQGAQLHEPPQYKTHLATATDPAFCASALGEROGLEED 364
 DB 305 TSDLVMLPGMAQNGVYLVGLQGAQLHEPPQYKTHLATATDPAFCASALGEROGLEED 364

```

Db 305 TSDVMTLPNIAVGTVLGFLQRCALPREFSHYETHLADPAFCASL-NAELDEBH 363
QY 365 VVGORCPQDCITLQNVASGL-----NHQTSVYAAYSVAAQALHNTLQCNASCP 416
Db 364 VMGQRCPCDDIMLQNLSSGLQNLASQOLHQLFATYAAYSVAAQALHNTLQCNVSHCH 423
QY 417 AODPKPMQLENNYNTLFVHGGPLRFPSGSGNVDMEDLKLWQSGVPRLDHVGRENG 476
Db 424 VSEHLPQMLENNYNTLFVHGGPLRFPSGSGNVDMEDLKLWQSGVPRLDHVGRENG 483
QY 477 SLRTERLIRWHTSDNQRVSRCSROCGEGRVRRVKGFSHCYDCVDCDCAEAGSYRONPDI 536
Db 484 TLQLOQSKMYN--PQNVQPVSGCSROCKDQGVRRVKGFSHCYDCVDCDCAEAGSYRONPDI 541
QY 537 ACTPCGODEMSPERSTRCFRRRSRFLANGEPAYVLLLLLSLALGVLAAAGLFPVHHNS 596
Db 542 TCTPCNODMSPEKSTACLPKPKFLANGEPAYVLLLLLSLALGVLAAAGLFPVHHNS 601
QY 597 PLYVAGSGPLACFGLVCLGVLCLSVLFPGQSPARCIAQOPLSLPTGCLSTFLQAA 656
Db 602 PLYVAGSGSGQFCGLICGLICGLISVLPFGRRSSASCIAQOPMHLPLTGCLSTFLQAA 661
QY 657 EIFVESELPLSMADRLSGCLGPGMAVLVLLAMLEVALCTWYLAEPPEVYTDHMLPT 716
Db 662 EIFVESELPLSMANMLCSYLRGLMAMLVLLATFEVALCAMYLTAFPEVYTDHMLPT 721
QY 717 EALVHCRRSVWSGCLATNATNATLAFCLGTFVLRSGPGCYNRARGITFMAIYFTWV 776
Db 722 EYLEHCHRVSWVSGLVITNMAIAFLCGLTFLVQSGPGRNRRARGITFMAIYFTWV 781
QY 777 SEVPLLANVQVAVPAGVOMGALLCVLGIILAFHLPRCYLMBROGLTPEFFLGSGGPD 836
Db 782 SEVPLLANVQVAVPAGVOMGALLCVLGIILAFHLPRCYLMBROGLTPEFFLGSGGPD 841
QY 837 AOGONDGNTGNOGKHE 852
Db 842 AADENSG--GGEAAOE 855

RESULT 6
Q923KO PRELIMINARY; PRT; 858 AA.
ID Q923KO
AC Q923KO;
DT 01-DEC-2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Sweet taste receptor TIR3.
GN TASIR3 OR SAC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090.
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SWR/J;
RA Nelson G., Hoon M.A., Chandrasekar J., Zhang Y., Ryba N.J.P.,
RA Zuker C.S.;
RT "Mammalian Sweet Taste Receptors.";
RL Cell 0:0-0(2001).
DR EMBL: AY032621; AAK51602.1; -.
DR MGD: MGI:1933547; Tasir3.
DR InterPro: IPR001828; ANF_receptor.
DR InterPro: IPR000337; GPCR_Mgr.
DR Pfam: PF00003; 7tm_3; 1.
DR Pfam: PF01094; ANF_receptor; 1.
DR PROSITE: PS00980; G_PROTEIN_RECCEP_F3_2; UNKNOWN.1.
DR PROSITE: PS50259; G_PROTEIN_RECCEP_F3_4; 1.
KW Receptor.
SQ SEQUENCE 858 AA; 94530 MW; 686A7E524267796A CRC64;

Query Match 71.2%; Score 3221.5; DB 11; Length 858;
Best Local Similarity 72.4%; Pred. No. 8e-250;
Matches 620; Conservative 67; Mismatches 156; Indels 13; Gaps 4;

```

```

QY 5 AVLGSLMALHPGICAPLCLISQOLRMKGDVYLGSLFGLAEBAEAGLRSTRPSSPVCTR 64
Db 5 AIMGSLIAAFELGAGELISQOFKAQGDYILGLFGLGSTEETATLQRPNSILCNR 64
QY 65 FSSNGLIWAALMAKVAEEINNKSDLLPGLRGYDLPDCESEPVYAMKPSLMEFAAGSRD 124
Db 65 FSPGLGLFAMAMKVAEEINNKSSALLPGLRGYDLPDCESEPVYAMKPSLMEFAAGSRD 124
QY 125 IAAYCNYTQYQPRVLAIVGPHSSSELAVTGRFSEFFLMPQVSYGASMBELLSARETFPSEF 184
Db 125 IAAYCNYTQYQPRVLAIVGPHSSSELATLTKFSEFFLMPQVSYGASMBELLSARETFPSEF 184
QY 185 RTVPSDRVQLTAAAEELLQEFQGNWYAAAGSDDEYGRQSLSTFSALAAARGICIAHEGLVP 244
Db 185 RTVPSDRVQLTAAAEELLQEFQGNWYAAAGSDDEYGRQSLSTFSALAAARGICIAHEGLVP 244
QY 245 LPRADDSRLGKQVDLHVNQSVQVYLLFASVHAHLAFNYSSIRLSPKRWYVASEAML 304
Db 245 QHDTSGQDLGKVDLVCVQNSKVQVYLLFASARAVISLFSYSIHGLSPKRWYVASEAML 304
QY 305 TSDVMTLPNIAVGTVLGFLQRCALPREFSHYETHLADPAFCASL-NAELDEBH 363
Db 305 TSDVMTLPNIAVGTVLGFLQRCALPREFSHYETHLADPAFCASL-NAELDEBH 363
QY 365 VVGORCPQDCITLQNVASGL-----NHQTSVYAAYSVAAQALHNTLQCNASCP 416
Db 364 VMGQRCPCDDIMLQNLSSGLQNLASQOLHQLFATYAAYSVAAQALHNTLQCNVSHCH 423
QY 417 AODPKPMQLENNYNTLFVHGGPLRFPSGSGNVDMEDLKLWQSGVPRLDHVGRENG 476
Db 424 VSEHLPQMLENNYNTLFVHGGPLRFPSGSGNVDMEDLKLWQSGVPRLDHVGRENG 483
QY 477 SLRTERLIRWHTSDNQRVSRCSROCGEGRVRRVKGFSHCYDCVDCDCAEAGSYRONPDI 536
Db 484 TLQLOQSKMYN--PQNVQPVSGCSROCKDQGVRRVKGFSHCYDCVDCDCAEAGSYRONPDI 541
QY 537 ACTPCGODEMSPERSTRCFRRRSRFLANGEPAYVLLLLLSLALGVLAAAGLFPVHHNS 596
Db 542 TCTPCNODMSPEKSTACLPKPKFLANGEPAYVLLLLLSLALGVLAAAGLFPVHHNS 601
QY 597 PLYVAGSGPLACFGLVCLGVLCLSVLFPGQSPARCIAQOPLSLPTGCLSTFLQAA 656
Db 602 PLYVAGSGSGQFCGLICGLICGLISVLPFGRRSSASCIAQOPMHLPLTGCLSTFLQAA 661
QY 657 EIFVESELPLSMADRLSGCLGPGMAVLVLLAMLEVALCTWYLAEPPEVYTDHMLPT 716
Db 662 EIFVESELPLSMANMLCSYLRGLMAMLVLLATFEVALCAMYLTAFPEVYTDHMLPT 721
QY 717 EALVHCRRSVWSGCLATNATNATLAFCLGTFVLRSGPGCYNRARGITFMAIYFTWV 776
Db 722 EYLEHCHRVSWVSGLVITNMAIAFLCGLTFLVQSGPGRNRRARGITFMAIYFTWV 781
QY 777 SEVPLLANVQVAVPAGVOMGALLCVLGIILAFHLPRCYLMBROGLTPEFFLGSGGPD 836
Db 782 SEVPLLANVQVAVPAGVOMGALLCVLGIILAFHLPRCYLMBROGLTPEFFLGSGGPD 841
QY 837 AOGONDGNTGNOGKHE 852
Db 842 AADENSG--GGEAAOE 855

RESULT 7
Q920R8 PRELIMINARY; PRT; 840 AA.
ID Q920R8
AC Q920R8;
DT 01-MAY-1999 (Tremblrel. 10, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Putative taste receptor Trl (Fragment).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

```

OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MISTAR;
 RX MEDLINE=99159821; PubMed=10052456;
 RA Hoon M.A., Adler E., Lindemeier J., Battey J.F., Ryba N.J.,
 RA Zuker C.S.;
 RT "Putative mammalian taste receptors: a class of taste-specific GPCRs
 RT with distinct topographic selectivity.";
 RL Cell 96:541-551(1999).
 DR EMBL: AF127389; AAD18069.1; -
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCR_MGR.
 DR PROSITE: PS50259; G_PROTEIN_RECPT_F3_4; 2.
 KW Receptor.
 FT NON_TER
 SQ SEQUENCE 840 AA; 93496 MW; 1FCBE7EFC6B845DB CRC64;

Query Match 26.7%; Score 1208.5; DB 11; Length 840.

Best Local Similarity 33.8%; Pred. No. 3e-88; Matches 291; Conservative 151; Mismatches 339; Indels 79; Gaps 18;

QY 12 WALLHPTGAPLCISQOL-----RMKGDYVLGGLFPL-GEAEAGLR 53
 DB 4 WA-----AHLLSLQLVYCWAFSCQRTSSPGFSLPGDFLAGLFLSHGDC---LQV 52
 QY 54 RTRSSPVCR---FSSNGLLMALAMKAVEINNKSDLPGLGLGDFPTCEPVVAM 110
 DB 53 RHRPLVTSQDPSDFNGHGHFLFOAMRFTVEINNSLLPNTLGLYELDYCVES-ANV 111
 QY 111 KPSIMFLAKAGSRDIAYCNYTOYQPRYLAVIGPHSESLAMVTKFSEFLMPQVSGAS 170
 DB 112 YATIRVALGQPRHIEIOKDLRNHSSKVVAFIGPDNDTHATTAALGFLMPPLVYSAS 171
 QY 171 MELISARETPSEFRIVPSDRVQLTAAAEILOEFGMMVVAALGSDDEXGROGLSIFSALA 230
 DB 172 SVVLSAKRKPSFLRTVPSDRHGVEMVQQLDSFGMWISLIGSYGQGLGVQALEBIA 231
 QY 231 AARGICIAHEGLVPLPRADSRIG--KYQVDLHVNOSSVQVVLVFAVNAHALFNKSI 288
 DB 232 VPRICVAFKQIVLP---SARVSDPRQSMOHLAQRFTVVVFSNRHLARVEFSNV 287
 QY 289 SSRSPVYVWVAEAMLTSDLVMGRLPGMAOMQTVLGFLOGA---LHEFPQ-YVKTHTAL 344
 DB 288 LANLTGKVVWVAEEDMAISTYITSTVGIGTIGLVAVVQQRQVGLKEFESESYRAVTA 347
 QY 345 AT---DPAFCGALGERDGLDEVDVVGORPCQDCITLQNVN--AGLNHHOTFSVYAAYS 399
 DB 348 PSACPESGWS---TNOLCRECHFTTRNMPTLGAFSMSAAYRVEAYVA 394
 QY 400 VAOALNHTLQCNMAGCPADVPKVMOLLENNYNLTFFVGGRLPFRDSSGNDMEYDKLM 459
 DB 395 VAGLHOLLGCTSTIC-SRGEVYPMQLDQYKVFLLHENTVAFDDNGDGLGYDDITAM 453
 QY 460 VMQGSVPLRHDVGRFNS---LATERLKIRWHTSDNOKPVSRCSROCEQGVRRYKFS 516
 DB 454 DMNRPETFEIGSASISPHLDINKKIQHGNKNQVPVSVCTDCLAGHNRVVS 513
 QY 517 CCHVCVCEAGSYRQNDDIACFTGODEMSPSTRCFRRRSRFLAMGERAVALLLL 576
 DB 514 CCEPCVCEAGTFLNMEELHICQPCGTEEMAPKSESTCFPRVTEFLAMHEPISLITLAN 573
 QY 577 SLALGIVLAALGLFVHNRDPLVQASGRLACFGLVCLGLVCLSVLFPQSPARCLAQ 636
 DB 574 TLLLLLVGTRAGLFAHFHPPVNASAGRICETMLGSLVAGSCFYSFEGPVPACILR 633
 QY 637 OPTSLPLTGLSTFLQAAEIFV---ESELPL---SMADRLSGCLRGPAAMLVLLAM 689
 DB 634 OPTSLGFAIFLGLTIRSPOLVLIIFKSTKVPFFYRTWA-----QNHGAGLFVIYSS 686

QY 690 LVEVALCTWYLVAEPPEVNTDWMHLPTREALVHCRTSRWSVFGLAHATNATLAFCLGTF 749
 DB 687 TVHLILCTWLVMTPTPTREYQRFPHLVILECTEVNSVGLAFTHNLLISTFVCSY 746
 QY 750 LVRSQPCYRARRALPTAMLAFTITWSFVPLANQVLRPAVQKALLLCVIGILAA 809
 DB 747 LGKELPENYENAKCVTSLLNFVSWIAFTMASTYQGSYLPVAVNLAGLTTLSGFSY 806
 QY 810 HLPKCYILMRQPGDTEPEEF 829
 DB 807 FLPKCYIILCRPELNTNEHF 826

RESULT 8

ID 099PG5 PRELIMINARY; PRT; 842 AA.
 AC 099PG5;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Putative sweet taste receptor TIR1 (Fragment).
 GN TASI1R OR GPR70.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129P3/J;
 RX MEDLINE=21030739; PubMed=11178737;
 RA Li X., Inoue M., Reed D.R., Hugue T., Puchalski R.B., Tordoff M.G.,
 RA Nimolaya Y., Beauchamp G.K., Bachmanov A.A.;
 RT "High-resolution genetic mapping of the saccharin preference locus
 RT (Sac) and the putative sweet taste receptor (Tir1) gene (Gpr70) to
 RT mouse distal chromosome 4.";
 RL Mamm. Genome 12:13-16(2001).
 DR EMBL: AF301162; AAK07092.1; -
 DR MGD: MGI:1927505; Tasi1r.
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000345; Cytc_heme_bind.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCR_MGR.
 DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
 DR PROSITE: PS50259; G_PROTEIN_RECPT_F3_4; 1.
 KW Receptor.
 FT NON_TER
 SQ SEQUENCE 842 AA; 93428 MW; 39739A2FF482D33F CRC64;

Query Match 26.5%; Score 1199.5; DB 11; Length 842;
 Best Local Similarity 33.2%; Pred. No. 1.6e-87;
 Matches 285; Conservative 159; Mismatches 340; Indels 75; Gaps 17;

QY 12 WALLHPTGAPLCISQOL-----RMKGDYVLGGLFPLGEAEAGLR 52
 DB 4 WA-----AHLLSLQLVYCWAFSCQRTSSPGFSLPGDFLAGLFLSHGDC---HADCLQ 53
 QY 53 SRTSPSPVCR---FSSNGLLMALAMKAVEINNKSDLPGLGLGDFPTCEPVVAM 110
 DB 54 VHRPLVTSQDPSDFNGHGHFLFOAMRFTVEINNSLLPNTLGLYELDYCVES 113
 QY 108 VAMPKSLMFLAKAGSRDIAYCNYTOYQPRYLAVIGPHSESLAMVTKFSEFLMPQVSY 167
 DB 114 YV---TLRVLAQOGTGHLEQRLRNHSSKVVAFIGPDNDTHATTAALSPFLMPVSY 170
 QY 168 GASMLISARETPSEFRIVPSDRVQLTAAAEILOEFGMMVVAALGSDDEXGROGLSIFS 227
 DB 171 EASSVYILSGRKRPSPFLRTIPSDKYQEVYVRLQDSFGMWISLVGSGYQGLGVQALE 230
 QY 228 ALAARGICIAHEGLVPL-PRADSRIGKQVQVDLHVNOSSVQVVLVFAVNAHALFNK 286
 DB 231 ELATPRGICVAFKQIVLPVLSAQAGDPRN---QRMLRLARAKTYVVVFSNRHLARVEFS 287

[illegible]


```

DE Taste receptor T1R1.
GN TAS1R1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN
[1]
SEQUENCE FROM N.A.
RA Nelson G., Hoon M.A., Chandrasekar J., Zhang Y., Ryba N.J.P.,
RA Zuker C.S.;
RT "Mammalian Sweet Taste Receptors.";
RL Cell 0:0-0(2001).
DR EMBL: AY032622; AKS1603.1; -.
DR MGD: MGI:1927505; Tas1r1.
DR InterPro: IPR001828; ANF_receptor.
DR InterPro: IPR000345; Cytc_heme_bind.
DR InterPro: IPR000337; GPCR_Mgr.
DR Pfam: PF00003; 7tm_3; 1.
DR Pfam: PF01094; ANF_receptor; 1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE: PS50259; G_PROTEIN_RECP_F3_4; 1.
KM Receptor.
SQ SEQUENCE 842 AA; 93455 MW; D4DCE90959E991A CRC64;

Query Match 26.3%; Score 1188.5; DB 11; Length 842;
Best Local Similarity 33.1%; Pred. No. 1.2e-86;
Matches 284; Conservative 159; Mismatches 341; Indels 75; Gaps 17;

OY 12 WALHPTGAPICLSOOL-----RMKGDYVIGLFPLEGEAEAGLR 52
DB 4 WA-----AHLLISQLAAVAYCAFSQORTSSPGFSLPEDFLAGLFSL---HADCLQ 53
OY 53 SRRPSSPVCTR---FSSNGILMALAMKAVEEINNKSDLLPGRLGIDLEDTCSER--V 107
DB 54 VRHRPLVTSQDRSDSFNGHGHLEFQAMRFYEEINNTALLPNTITIGELAYDVCSESSNV 113
OY 108 VAMKPSLFLAKAGSRDIAATCANTQYQPRVLAIVGHSSELAVYTGKFSFLLMPQVSY 167
DB 114 YA---TLRVPAOQCTGHEMQRDRRNHSSKVVALLGPDNTDHAVTALLSPFLMPVSY 170
OY 168 GASMEILSARETFPSFRTPVSDRVOLTAALLOEGMMNVAALGSDDEXGROGLSIFS 227
DB 171 EASSVILSGKKRFPSELTITPSDKYQVEVYRLQSGFWWISLVGSGDGLQGLVQALE 230
OY 228 AALAAAGICIAHEGLVPL-PRADSRSLGKYQVLAHOVQSSQVVLLEFASVHAHALENY 286
DB 231 ELATPRGICVAFKDVVPLSAQADPRM---QRMALRLARATVYVVESENHLLGVFERS 287
OY 287 SISRLSPKRWVVAEAMLTSDLVNGLPGMAQMGTVLGLQAGDLHEPQYVKTHL-ALA 345
DB 288 VVLANLIGKVIASEDMAISTYITNVPGIOIGITVLGAIAIOQRPVGLKEFEESYQAVM 347
OY 346 TDPAFV---SALGREGGLEEDVVGORCPQDCITLQNV--AGLNHQRTSVAAVSY 400
DB 348 GAPTCTEGSGMG-----TNQLCRCHATTNNMPLGAFMSAAVNYEAYTAV 397
OY 401 AQAHLNLTQCNASGCPADPYKPMQLLENMNTLTFHVGGLEPLRDSGNDVMEDELKLV 460
DB 398 AHGHQLIGCTSGTC-ARGPYVPMQLQIYKVFLLHKKVAFDDKDDPLGYDIIAMD 456
OY 461 WQGVNPLHDVGRNGS---LTERLAKIRWHTSDNOKPVSCSQCOCQGVRRYKGRHSC 517
DB 457 WNGEMTFEYVYASLSLSPVHLIDINKTKITOMHGKNNQVAVSVCTRDCEGHRLVWGSNHC 516
OY 518 CYDVPDEAGSYRONPDIACTPCGODEMSPERSTRCFRRRSRLAMCEPVLILLILLS 577
DB 517 CFEBMPCEAGFLNTSELHTQPCGTEMAPESSACFSRIVEELGWEPISLVLANT 576
OY 578 IALGLVLAALGLFVHNHSDPLVQASGGLACFLGVLVCLVLSLTPGQSPARACQAO 637
DB 577 LLLLLLIGTAGLFWRLHTFPVRSAGRLCTLMGLSIVAGSCSLXSPFGKTYVACILRQ 636
OY 638 PLSHLPLTGLCLSTLFLQAAELFV---ESELPL--SWADRLSGCLRGPMAMLVVLLAML 690

```

```

DB 637 PLFSLGPAIFLSCLTIRSFOLVIFKFKSTKVPTEYFHTWA-----ONHGAGIFVIYST 689
OY 691 VEVALCTWYLVAFPEPVVTOHMLPTBALVHCRRMSVSGCLANATATLAFCLCTFL 750
DB 690 VHEFLCTWLAAMTPRTRTEORFPHLVILECEVNSVGLFAVANHILISTFVCSYL 749
OY 751 VRSQPGCYNRARGLTFPMLAVFTWVSFVPLIANVQVLRPAVOMGALLCVLGITLAFH 810
DB 750 GKELPENYNENKCVTFSLLLHFWSIWIAFTMSSLYQGSYLPVAVVLAGLMTLSGSGGYF 809
OY 811 LPRCYLLMRQGLNTPPEF 829
DB 810 LPKCYVILCRPELNTNTEHF 828

RESULT 11
OY99PG6 PRELIMINARY; PRT; 842 AA.
AC O99PG6;
DT 01-JUN-2001 (TREMURel. 17, Created)
DT 01-JUN-2001 (TREMURel. 17, Last sequence update)
DT 01-JUN-2002 (TREMURel. 21, Last annotation update)
DE Putative sweet taste receptor T1R1 (Fragment).
GN TAS1R1 OR GPR70.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN
[1]
SEQUENCE FROM N.A.
RC STRAIN=C57BL/6BYJ;
RX MDLINE=21030739; PubMed=11178737;
RA Li X., Inoue M., Reed D.R., Hugue T., Puchalski R.B., Tordoff M.G.,
RA Nishimura Y., Beauchamp G.K., Bachmanov A.A.;
RT "High-resolution genetic mapping of the saccharin preference locus
RT (sac) and the putative sweet taste receptor (T1R1) gene (Gpr70) to
RT mouse distal chromosome 4.";
RL Mamm. Genome 12:13-16(2001).
DR EMBL: AF301161; AK07091.1; -.
DR MGD: MGI:1927505; Tas1r1.
DR InterPro: IPR001828; ANF_receptor.
DR InterPro: IPR000345; Cytc_heme_bind.
DR InterPro: IPR000337; GPCR_Mgr.
DR Pfam: PF00003; 7tm_3; 1.
DR Pfam: PF01094; ANF_receptor; 1.
DR PRINTS: PR00248; GPCRGR.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
DR PROSITE: PS50259; G_PROTEIN_RECP_F3_4; 1.
KM Receptor.
FT NON_TER
SQ SEQUENCE 842 AA; 93425 MW; 58826C43F5D352E CRC64;

Query Match 26.1%; Score 1180.5; DB 11; Length 842;
Best Local Similarity 32.9%; Pred. No. 5.3e-86;
Matches 283; Conservative 159; Mismatches 342; Indels 75; Gaps 17;

OY 12 WALHPTGAPICLSOOL-----RMKGDYVIGLFPLEGEAEAGLR 52
DB 4 WA-----AHLLISQLAAVAYCAFSQORTSSPGFSLPEDFLAGLFSL---HADCLQ 53
OY 53 SRRPSSPVCTR---FSSNGILMALAMKAVEEINNKSDLLPGRLGIDLEDTCSER--V 107
DB 54 VRHRPLVTSQDRSDSFNGHGHLEFQAMRFYEEINNTALLPNTITIGELAYDVCSESSNV 113
OY 108 VAMKPSLFLAKAGSRDIAATCANTQYQPRVLAIVGHSSELAVYTGKFSFLLMPQVSY 167
DB 114 YA---TLRVPAOQCTGHEMQRDRRNHSSKVVALLGPDNTDHAVTALLSPFLMPVSY 170
OY 168 GASMEILSARETFPSFRTPVSDRVOLTAALLOEGMMNVAALGSDDEXGROGLSIFS 227
DB 171 EASSVILSGKKRFPSELTITPSDKYQVEVYRLQSGFWWISLVGSGDGLQGLVQALE 230

```

```

QY 228 ALAARGICIAHEGLVPL-PRADDSRLGVODVLHGVNOSVYVLLFASVAAHAALPNY 286
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 221 ELATPRGICVAFKNVYPLSAQADPRM---QRMALRLARTYVYVSNRHLGVFRS 287
QY 287 SISRSLPRVAVASFWLTSVLVGLPGMAQWGLVGLQGAOLHEEPQYVTKHL-ALA 345
  :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 288 VVLALVTGKVIASBDMAISTNIPNVSGIIGIVLVAIGVIAOORVPGIKKEFEESVQAVM 347
QY 346 TDRPAC---SALGEROGLLEEDVYQORPCDCITLQVNS--AGLNHQTSTVYAAVSY 400
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 348 GAPRCPSPGSMGCG-----TNQICRECAFTTMMNPELGAFMSMAAYNAVYAVAY 397
QY 401 AQAALNTLOCNAGSGCPADDPVYKPMOLLENNYMLTFHVGLPLRFDSSGNVMEYDLKLMV 460
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 398 AHGLHQLGCTSGTC-ANGPHNPMOLLOQIYKVNFLKRYAFDPKDDPLGYDITIAM 456
QY 461 WOGSPRLHDVGRFNGS---LTERELKTRMHTSDNOKPVSRCSROCGOVRVYKGFHSC 517
  :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 457 WNGPEMTEFVIGSASLSPVHLDINKTKIOMHGKKNQVPSVCTRDCLSGHHLVWGSHHC 516
QY 518 CYDDVDCAGSGVROMPDILACTFCGODEMSPERSTRCFRRSRFLANGEPVALLLLLS 577
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 517 CFECPCPAGFTPLNTSELHTQPCGTEEMAPBSACFSRTVEFLGHEPISLVLLANT 576
QY 578 LALGIVLALGLFVHRDPSVVOASGPGVPLACGLVCLVCLSVLLFPGQSPARCIAQO 637
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 577 LILLILLITAGIFAMRLRTFPVRSAGRLCPLMLGSLVAGSCSLXSFGRKPRVPCILRQ 636
QY 638 PLSLHPLTGLCTSTFLQAAEIFV---BSELPL---SWADRLSGCLRGPWMLVYLAML 690
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 637 PLFSGIFAFISCLTIRSFQVILIRFSTKVTFYHTWA-----QNHGAGIEFYVST 689
QY 691 VEVALCTWYVLAAPREVYTDHMLPTEALVHCRTSWSPFGLAHNTATLAFCLGLFPL 750
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 690 VHLFCITLWLMWTPRPTREYQRPVHLVLCCTEVNSGVFLVAFNHLILISTEVCSTL 749
QY 751 VRSOPGCYNRARGLTFAMLAFTVWSEFVPLIANVQVVLRAVOMGALLLCVILIAAFH 810
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 750 GKELPENYENKCVTFSLILHFVSINIAFTMSIYGSYLRAVNVLAGLATLSGFSGYF 809
QY 811 LPRCLLMROGCLNTPPEFF 829
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 810 LPRCYVILCRPELNTNTEHF 828

RESULT 12
Q87D39 PRELIMINARY; PRT: 763 AA.
AC Q87D39;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Gml48 form B.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE-21888635; PubMed-11891061;
RA Makalowska I., Sood R., Faruque M.U., Hu P., Robbins C.M.,
RA Edlings E.M., Mestre J.D., Baxevanis A.D., Carpten J.D.;
RT "Identification of six novel genes by experimental validation of
RT Genemachina predicted genes.";
RL Gene 284:203-213(2002).
DR EMBL; AF367618; AAL91359.1;
SQ SEQUENCE 763 AA; 84427 MW; 593616575D6BD17D CRC64;

Query March 23.8%; Score 1077.5; DB 4; Length 763;
Best Local Similarity 33.0%; Pred. No. 8.5e-78;
Matches 256; Conservative 140; Mismatches 332; Indels 47; Gaps 14;
QY 76 KMAVEEINNSDILLGLRGLYDLFDTCSEPVVAMKPSLMLFLAKGSRDIAAYCVYQYQ 135

```

```

Db 1 MRLGEEINNSTALLPNTLTIGQLYDVCSDS-ANYATLRLVLSLPGQHIELQGLLHYS 59
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 136 PRVIAVIGPSSSELMATVTKFFSEFLMPQVSYGASMEILSARETFSEFFRVPSDRVOLT 195
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 60 PTVIAVIGPDSSTNRATPAALSLPPLVPMISAAASSEFLSVKQYPSLRTIPNDKYOVE 119
QY 196 AAELGDFGNWVAALSDDDEYGRGLSIFSAALAAAGICIAHEGLVPL-PRADDSRLG 254
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 120 TWVLLQKFGWTWSTLVSSSDYQGLQVLALENQATGGICAFEDIMFSAQVDEK-N 178
QY 255 KYODVLHGVNOSVYVLLFASVAAHAALFNYSISSRSPKVVYVASEAMLTSDLVMLGPG 314
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 179 --QCLMRLAQAATVYVYVSSRQIARVFESEVLTNLTKGKWASEMALSRHITGVP 236
QY 315 MAQGTVLGF-LQGA--QLHEFPQYVTKHALATDP---AFSGALGEROGLLEEDVYG 367
  :||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 237 IORIGMVIGVAIQKRAVAGLKAFEAVARADKXAPRCPCHKSGMSS-----N 283
QY 368 QRCPCDCI---TLQNVASGLNHQTFESVYAAVYVQAQALHNTLOCNAGSGCPADDPVYK 424
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 284 QLCRECAFMATMPEKLA-FSMSATYNAVYAVAHGLHQLG-C-ASGACSRGRTV 341
QY 425 QLENNYMLTFHVGLPLRFDSSGNVMEYDLKLMWOGSPVRLHDVGRFNGS--LRT 481
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 342 QLBQIHKVHFLHNDVYAFNDNRDPLSSYNIAIDMNGPKTFVYVLSSTWSPYQMLIN 401
QY 482 RIKTRMHTSDNOKPVSRCSROCGOVRVYKGFHSCCYDVCDEAGSYROMPDILACTFC 541
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 402 EFKIOMHGKKNQVPSVCSLCEGHQRYVYVTHHCCEVCAGATFLANKSDTLRCOPC 461
QY 542 GODEMSPERSTRCFRRSRFLANGEPVALLLLLSLALGIVLALGLFVHRDPSVVOA 601
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 462 KEEMAPBSQTCFRTVYFLALREHTSMVLLANTLILLITAGIFAMWLDTPVRS 521
QY 602 SGCPVPLACGLVCLVCLSVLLFPGQSPARCIAQOPLSHLPLTGLTFLQAAEIFV- 660
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 522 AGRCICFLMTGLSAGSSIVGFEXEPRPRCLRLQALFALGFTIFLSCLTVRSQQLII 581
QY 661 ---BSELPL---SWADRLSGCLRGPWMLVYLAMLVEVALCTWYVLAAPREVYTDHML 714
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 582 KFFSTKVPTFYHAM-----YONHGAGLFWISSAOLLILCTLWLVMTPLPAREYORF 634
QY 715 PTEALVHCRTSWSPFGLAHNTATLAFCLGFTLVRSOPGCYNRARGLTFAMLAFT 774
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 635 PLVWLECTETNSLGFILAFYVNLISAFACSTLGGDLDPENTYEAACVFTSLFNYS 694
QY 775 WVSFVPLIANVQVVLRAVOMGALLLCVILIAAFHPRCYLMLRPOGLNTPPEFF 829
  |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
Db 695 WIAFFTASVYDGKYLPAANMMAGLSLSSGFGYFLPCYVILCRPDLNSTEHF 749

RESULT 13
Q90WL6 PRELIMINARY; PRT: 940 AA.
AC Q90WL6;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE Extracellular calcium sensing receptor.
GN CASR.
OS Sparus aurata (Gilthead sea bream).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorpha; Acanthopterygii; Percomorpha; Perciformes; Percoidae;
OC Sparidae; Sparus.
OX NCBI_TaxID=8175;
RN [1]
RP SEQUENCE FROM N.A.
RA Clark M.S.;
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
SQ SEQUENCE FROM N.A.

```

RA Flanagan J.A.;
 RT "Calcium sensing receptor: cDNA sequence and mRNA expression in a
 marine teleost."
 RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AJ289717; CAC41352.1;
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PROSITE: PS00979; G_PROTEIN_RECIP_F3_1; UNKNOWN_1.
 DR PROSITE: PS50259; G_PROTEIN_RECIP_F3_4; 1.
 KM Receptor.
 SQ SEQUENCE 940 AA; 105455 MW; 19C3CEDC75A0E3F9 CRC64;

Query Match 23.4%; Score 1060; DB 13; Length 940;
 Best Local Similarity 30.4%; Pred. No. 2.9e-76;
 Matches 266; Conservative 150; Mismatches 341; Indels 118; Gaps 26;

QY 27 QQLRMKGDVYGLFPLGEAEAGLSR-----TRSSPVCTRFSSNGLLMALMKMAVE 81
 DB 24 QRAQMTGDIILGGLFPI-----HFGVASKDDDLAARPESSOCVRENFRGFRMLQAMIFAID 79

QY 82 EINNKSLLPLGLRLGYLDFPTCSEPVVAMKPSLMFLA--KAGSRDIAAYCNYQYOPRVL 139
 DB 80 EINNSSLPLPRLTYGRIFPTCNTVSKALEXTLSFVAQNKIDSLNLFDEFCNCTDHIPATI 139

QY 140 AVIGPHSSELAMWTGKFFSFPLMPQVSYGASMEILSARETFPFRTVPSDRVOLTAAE 199
 DB 140 AVVGAAGSASTAVANLLGLFPIQISYASSSRLSNKQKSMRMLPTDEYQATMAD 199

QY 200 LLOFGNWAALGSDDEYGRQGLSTISALAANGICLAHEGLVPLPRADSRIGKQDV 259
 DB 200 IIEFGNWWASAVASDDYGRPVGEKREKEMEEDICIHNLISQYFEDH---EIQAL 255

QY 260 LHQVNOSSQVVLLEFASVHAHAALFNYSISRLSPKTVVASEAMLT-----DLVM 310
 DB 256 ADRIENSTAKYIVFAFASGPDIPELIKEMVRNITDRIMLASEAMSSSLAKPEYLDVYA 315

QY 311 GLPQMA-OMGTVLGFLRGAGLHPEPQYVKT-----HLALATDPAFCS 352
 DB 316 GTIGFALKAGHIPIGF-----REFLOQVQPKDSDHNEFVREMEETFNCLIEDSP--- 364

QY 353 ALGRREGGL-----EEDVVGQRPCDCITLQNVASAGLNHQTFSYAAYVAQA 403
 DB 365 RLQSENGSTSFRLCTGEEDITSETPYLDYTHLR-----ISYVVAAYTSAQA 415

QY 404 LHNFLQNA-----SGCPADQVPRKPMOLLENNYNTLF--HYGGLPLRPDSSGVDMEYD 455
 DB 416 LQDILCTPQOGLFANNSCADIKKEMAKQVYKQRLHLYTNSMGEKMHFDENADLAANTT 475

QY 456 LKLM---VMQSVPRLDHVGFRN-----GSLRTERLKIHWTSDNQPVSRCSROCEG 506
 DB 476 LHMHRSAEDGSYV--FEVGYNNHAKRGKLFIDNTKILMNGFSSEVPSCNCEDEPG 534

QY 507 QVRR--VAGFHSCTCYDCVCEAGSTRONPDIACTFCGODEMSPKSTRCFRRSRFLAMG 565
 DB 535 TRRGIDISMTCTCECTECSCGEYSYHKDASVCTKCPNNSNNGNHTFCFELKEITFLSWT 594

QY 566 EPAVLILLLLSLALGLVLA--LGLFVHHRDSPLYOASGGPLA---CFGYVCGIWL- 619
 DB 595 EP--FGIALAICAVLAGVYVLAFAVGVFRFRNPTIVAKTNRELYVLLFSLIC---CFG 648

QY 620 SVLLPQSPARCLAQOPLSHLPLTCLSTLFLQAAEIFV--ESELPLSWADRLSGCLR 677
 DB 649 SLIFITQOPQDMQRLNPAFGISFVLCISGILVKTNRVLLVFAKIPTS-----LH 700

QY 678 GPM-----AMLVILLAMLEVALCTWYLVAPPEVVTDMHMLPTALVHCRTSWVSFGL 732
 DB 701 RKMWGLNLOGLFVAVLCTFVQVMICVWVLYNAPSSYNN--HDIDELITFCNEGVSMAIGF 759

QY 733 AHATNATLAFCLGTFVLSQPCCYNRARGLTFAMLAFTITWVSFPLANQVVLARA 792
 DB 760 LIGTYCLLAICFPFAKSKRLPENFTFAKFTFSMLIFLWISFLPAVFTYTGKFEVA 819

QY 793 VOMGALLCYGIIAFAHLPRCYLLMRQPLNTP 827
 DB 820 VEAIALIASSFGMLACIFFKFKVYILLKPRSRNTLE 854

RESULT 14
 ID 073635 PRELIMINARY; PRT; 940 AA.
 AC 073635;
 DT 01-AUG-1998 (TEMBLrel. 07, Created)
 DT 01-AUG-1998 (TEMBLrel. 07, Last sequence update)
 DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
 DE Calcium2+ sensing receptor.
 OS Fugu rubripes (Japanese pufferfish) (Takifugu rubripes).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 OC Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 OC Tetraodontidae; Takifugu.
 OX NCBI_TaxID=31033;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=9826788; PubMed=9560249;
 RA Naito T., Saito Y., Yamamoto J., Nozaki Y., Tomura K., Hazama M.,
 RA Nakashishi S., Brenner S.;
 RT "Putative pheromone receptors related to the Ca2+-sensing receptor in
 fugu.";
 RL Proc. Natl. Acad. Sci. U.S.A. 95:5178-5181(1998).
 DR EMBL: AB008857; BAA26122.1; -;
 DR InterPro: IPR001828; ANF_receptor.
 DR InterPro: IPR000337; GPCR_Mgr.
 DR InterPro: IPR000651; RasGEFN.
 DR Pfam: PF00003; 7tm_3; 1.
 DR Pfam: PF01094; ANF_receptor; 1.
 DR PRINTS: PR00248; GPCRMRG.
 DR PROSITE: PS00979; G_PROTEIN_RECIP_F3_1; 1.
 DR PROSITE: PS50259; G_PROTEIN_RECIP_F3_4; 1.
 KM Receptor.
 SQ SEQUENCE 940 AA; 105814 MW; 06DAB7803B6878B3 CRC64;

Query Match 23.2%; Score 1049; DB 13; Length 940;
 Best Local Similarity 29.5%; Pred. NO. 2.2e-75;
 Matches 235; Conservative 157; Mismatches 357; Indels 96; Gaps 21;

QY 26 SOQLRMKGDVYGLFPLGEAEAGLSR-----TRSSPVCTRFSSNGLLMALMKMAV 80
 DB 24 NQRAQMTGDIILGGLFPI-----HFGISSKDENLAAARESTRKYVFRNRRGRMQLAMVFAI 79

QY 81 EINNKSLLPLGLRLGYLDFPTCSEPVVAMKPSLMFLA--KAGSRDIAAYCNYQYOPRV 138
 DB 80 EINNSSLPLPRLTYGRIFPTCNTVSKALEXTLSFVAQNKIDSLNLFDEFCNCTDHIPAT 139

QY 139 LAVIGPHSSELAMWTGKFFSFPLMPQVSYGASMEILSARETFPFRTVPSDRVOLTAAE 198
 DB 140 IAVVGAAGSASTAVANLLGLFPIQISYASSSRLSNKQKSMRMLPTDEYQATMAD 199

QY 199 ELLOEFGMNWAALGSDDEYGRQGLSTISALAARGICIAHEGLVPLPRADSRIGKQDV 258
 DB 200 DVIEFPMNNVAVASDDYGRPGIEKFEKEMERDICHNLISQYFED---CEIKA 255

QY 259 VLVHNOSSQVVLLEFASVHAHAALFNYSISRLSPKTVVASEAMLTSDLVMLPGMAQ 318
 DB 256 LVRIENSTAKYIVFAFASGPDIPELIKEMVRNITDRIMLASEAMSSSLAKPEYLDV 315

QY 319 GTVIGFLQRAQL---HEFPQYVKT-----HLALATDPAFCSALGERE 358
 DB 316 EGTIGFLVAKGNLPGREFLQVQPKRGSNEVREFWEETFNCLIEDSP---RLQSE 371

QY 359 QG-----LEEDVVGQRPCDCITLQNVASAGLNHQTFSYAAYVAQAALHTLQ 409
 DB 372 NGSISFRLCTSEEDITSETPYLDYTHLR-----ISYVVAAYTSAIARLQDILS 422

QY 410 CNA-----SGCPADQVPRKPMOLLENNYNTLF--HYGGLPLRPDSSGVDMEYDLKM-- 459

```

Db 423 CTPHGGLFANNSCADIKKEMAMQVLFKQRLHNTYNSMGEKXHFEDNMAMENYTTINHR 482
Qy 460 -VMQGSVPRLDVG-----RPNGLRTERLKIRNHTSDNOKPVSRGSCQOEGVRR-V 511
Db 483 SAEGSV-AFREVGYYHMHARGAKLLIDNTRKMMNNAVSSEVPSCSEDEPGTRKGI 541
Qy 512 KGFHSCCDYDCEAGSYRONDDIACFFCGODEMSPERSRRCFRARRFLAMGEPAYLL 571
Db 542 DSMPTCCCECTECSDGEISDKKDSICTKCPNNSKSSNHTFCHLKEFLFAMSSEPGIA 601
Qy 572 LLLSLALGLVLAAGLGFVHHRDSPVLOASGGLACGLVCLVCL--SVLLPFGOPS 629
Db 602 LAICAVIGLVLLAFVAGVFRNTPYKASNRLEST--VLLSLICCFSSLLIFIGEPQ 659
Qy 630 PARCAQOPLSHLPITGLSTFLQAAETIV--SEELPLS-----WADRLSGCLRGPMW 682
Db 660 DMTGRRLRQPAIGISFVLCISCIIVKTRVLLVFBAKIPTSIRKRWGINTLQ-----F 711
Qy 683 LVVLLAMLEVALCTWYLVAFRPEVVDHMLPFEALVHCSTRSMVSFGLAHATNATLAF 742
Db 712 LVVLCITFVQVVICVWLYNAPSSYRN-HDIDELITFCBGSWVALGFLIGYCLIAA 770
Qy 743 LCLFGLTVRSQPCGYNRARGLTFAMLAFTTWVSFVPLANVQVLRPAVQMGALLICV 802
Db 771 ICEPFAFRSKLPENFTFAKFTFCMLIFFTWISFIPATYSTYKGVSAVEALAILASS 830
Qy 803 LGITAAFLPRCYLLMRQPLNTEP 827
Db 831 YGMALCIFENKRYITLFRKCRNTE 855

```

RESULT 15

```

ID 08TE23 PRELIMINARY; PRT: 839 AA.
AC 08TE23;
DT 01-JUN-2002 (TREMblrel, 21, Created)
DT 01-JUN-2002 (TREMblrel, 21, Last sequence update)
DT 01-JUN-2002 (TREMblrel, 21, Last annotation update)
DE Taste-specific G-protein coupled receptor T1R2.
GN T1R2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21927605; PubMed=11917125;
RA Li X., Staszewski L., Xu H., Durick K., Zoller M., Adler E.;
RT "Human receptors for sweet and umami taste."
RL Proc. Natl. Acad. Sci. U.S.A. 99:4692-4696(2002).
DR EMBL: AF458154; AAM12239.1; JOINED.
DR EMBL: AF458150; AAM12239.1; JOINED.
DR EMBL: AF458151; AAM12239.1; JOINED.
DR EMBL: AF458152; AAM12239.1; JOINED.
DR EMBL: AF458153; AAM12239.1; JOINED.
KW Receptor.
SO SEQUENCE 839 AA; 95072 MW; 8CED6F1503F05FDE CRC64;

```

Query Match 22.4%; Score 1014.5; DB 4; Length 839;
 Best Local Similarity 30.7%; Pred. No. 1.1e-72;
 Matches 264; Conservative 152; Mismatches 378; Indels 67; Gaps 22;

```

Qy 2 LGPAV-----LGLSMALLHETGAPLCISQQLRMKGDYVGLGFLPLGEAEAGLRSKR 56
Db 1 MGPRAKTTICSEFFLLMVAEPNDSFYL-----PGDYLLGLGLFSL--HANMKGIIVHLNF 53
Qy 57 PSSPYCTFRSSNGILMAL--AMKMAVEEINNKSDLLPGLRGLGYDLFDC--SEPVVAMKP 112
Db 54 IQVPCKEYEVKIVGINLMQAMRFAVEEINNDSSLLPGVLLGLEYIVDCYISNNV---QP 110
Qy 113 SLMLAKGSRDIAAYCNYTOYQPRVLAVIGPHSSELMAYTGKFFSFLLMPQVSYGASME 172

```

```

Db 111 VLYFLAHBDNL-LPQGEYSNYSIRVAVIGDENSESVMTVANFLSLFLPLQITYSALSD 169
Qy 173 LLSARETPSFRRYVSPRRVOLTAAAEILOEGCMWVAALSDDDEYQGQGISFALAAA 232
Db 170 ELRKVRPRALLRTPRSADHVEAMVQIMLHERMMIIVLVSSDIYGHNDQLGERVAR 229
Qy 233 RGICIAHGLVPLPRADSRGL---KVQDVLHQVNOSSVQVLLFAFVAHAHALFNYSI 288
Db 230 RDICIAFETLEPITLPNQMNTSEERQRLTYIDKIQOSTARVYVFSDDLTIHFENEVL 289
Qy 289 SSRSPKVVASAEAMTSDLVWGLPMAQMGTVLGLRQAGLHFE-----POYKTH 341
Db 290 RQNFQAVWIASSESAIDPVLHNLDELHGLGFLGITIQSVIPQFSEFEREMGPQAGPP 349
Qy 342 LALATDPAFCSALGEREGLEDVYGRCPQDCITLQ--NVSAGLNHQ--TTSYVAAYS 399
Db 350 LSRTSQSYTCN-----QEDCNCLNMTLSFNTLLRISGERVYVSYSAYYA 394
Qy 400 VAQALHNTLOCNAGCPADDPYKPMQLENNYNTLFHYGGLPLRPDSSQNVDMEDLKM 459
Db 395 VAHALHSLGCDKSTC--TKRVYYPQQLBEIKVFTLLDHOIFDPQGDVALHLEYQW 453
Qy 460 VMQGSVPRLDHVGRENGSLRTERLK----IRNHTSDNOKPVSRGSCQOEGVRRYKGFH 515
Db 454 QMDRSQNFQSVASYYPQOR--QLKNIDISHTVNTNTPMSCKRQSGQKKRPVGIH 511
Qy 516 SCCTYCVCEAGSY--RONDDIACFFCGODEMSPERSRRCFRARRFLAMGEPAYLL 574
Db 512 VCCFPCIDCLPQTFNHTDEDEYCOACPNNMYSQSESCFRQLVFLMHAPPIAVAL 571
Qy 575 LLSLALGLVLAAGLGFVHHRDSPVLOASGGLACGLVCLVCLSVL-LFPQSPARC 633
Db 572 LAAIGFLSTALIVLFMRHFORPIYRSAGPR--CLMLTLILVAAMVYVYGGPKVSGC 630
Qy 634 LAQOPLSHLPITGLSTFLQAAETIVESELPNSADRISGLR--GPM--AMLVILIA 688
Db 631 LCRQALFPLCFICISCIASIAVRSFOIVCAFKMASRFPRAYSYWRVQGPVSNAFITVLKM 690
Qy 689 MIVEVALCTWYLVAFRPEVVDHMLPFEALVHCSTRSMVSFGLAHATNATLAFICPLGT 748
Db 691 VIVVIGMLA---TGLSPTRTPDD--PKITTVSCNPNRNSLIENTSDDLISVGFSTA 746
Qy 749 FLVRSQPCGYNRARGLTFAMLAFTTWVSFVPLANVQVLRPAVQMGALLICVILIAA 808
Db 747 YMGKELPTNYNNAKFTILSMTEFTSSVSLCFMNASVGLVTVIYDLTVNLALASIG 806
Qy 809 FHLPRCYLLMRQPLNTEPEF 829
Db 807 YFGPRCYMLTFEPRNTPAYF 827

```

Search completed: May 19, 2003, 09:52:40
 Job time : 41.8189 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:48:40 ; Search time 14.6201 Seconds
(without alignments)
1714.646 Million cell updates/sec

Title: US-09-927-315-15

Perfect score: 4524

Sequence: 1 MGPAYGLSLMALHPGTG.....GPGDAQGQNDGNTGNQKHE 852

Scoring table:

BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 262574 seqs, 29422922 residues

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: Issued_Patents_AA:*
2: /cgn2_6/ptodata/1/1aa/5A.COMB.pep:*
3: /cgn2_6/ptodata/1/1aa/5B.COMB.pep:*
4: /cgn2_6/ptodata/1/1aa/6A.COMB.pep:*
5: /cgn2_6/ptodata/1/1aa/PCTUS.COMB.pep:*
6: /cgn2_6/ptodata/1/1aa/Backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1089.5	24.1	1059	4 US-09-134-513-2	Sequence 2, Appli
2	1068.5	23.6	1078	1 US-08-485-588-7	Sequence 7, Appli
3	1068.5	23.6	1078	1 US-08-484-565-7	Sequence 7, Appli
4	1068.5	23.6	1078	2 US-08-480-751-7	Sequence 7, Appli
5	1068.5	23.6	1078	2 US-08-943-986-7	Sequence 7, Appli
6	1068.5	23.6	1078	3 US-08-353-784-7	Sequence 7, Appli
7	1068.5	23.6	1078	4 US-08-484-719B-7	Sequence 7, Appli
8	1068.5	23.6	1078	4 US-08-484-159-7	Sequence 7, Appli
9	1060	23.4	1085	1 US-08-485-588-5	Sequence 5, Appli
10	1060	23.4	1085	1 US-08-484-565-5	Sequence 5, Appli
11	1060	23.4	1085	2 US-08-480-751-5	Sequence 5, Appli
12	1060	23.4	1085	2 US-08-943-986-5	Sequence 5, Appli
13	1060	23.4	1085	3 US-08-353-784-5	Sequence 5, Appli
14	1060	23.4	1085	3 US-08-484-719B-5	Sequence 5, Appli
15	1060	23.4	1085	4 US-08-484-159-5	Sequence 5, Appli
16	1057.5	23.4	1088	1 US-08-485-588-6	Sequence 6, Appli
17	1057.5	23.4	1088	1 US-08-484-565-6	Sequence 6, Appli
18	1057.5	23.4	1088	2 US-08-480-751-6	Sequence 6, Appli
19	1057.5	23.4	1088	2 US-08-943-986-6	Sequence 6, Appli
20	1057.5	23.4	1088	3 US-08-353-784-6	Sequence 6, Appli
21	1057.5	23.4	1088	3 US-08-484-719B-6	Sequence 6, Appli
22	1057.5	23.4	1088	4 US-08-484-159-6	Sequence 6, Appli
23	1056	23.3	1027	4 US-09-162-021B-2	Sequence 8, Appli
24	1054	23.3	1079	1 US-08-485-588-8	Sequence 8, Appli
25	1054	23.3	1079	1 US-08-484-565-8	Sequence 8, Appli
26	1054	23.3	1079	2 US-08-480-751-8	Sequence 8, Appli
27	1054	23.3	1079	2 US-08-943-986-8	Sequence 8, Appli

28	1054	23.3	1079	3 US-08-353-784-8	Sequence 8, Appli
29	1054	23.3	1079	3 US-08-484-719B-8	Sequence 8, Appli
30	1054	23.3	1079	4 US-08-484-159-8	Sequence 8, Appli
31	1009.5	22.3	877	4 US-09-619-353-2	Sequence 2, Appli
32	1002.5	22.2	843	4 US-09-361-631-1	Sequence 1, Appli
33	999.5	22.1	863	4 US-09-619-353-14	Sequence 2, Appli
34	974	21.5	863	4 US-09-619-353-14	Sequence 14, Appli
35	938.5	20.7	1219	2 US-08-687-289A-6	Sequence 6, Appli
36	839	18.5	835	4 US-09-619-353-7	Sequence 7, Appli
37	788.5	17.4	851	4 US-09-619-353-12	Sequence 12, Appli
38	787	17.4	856	4 US-09-619-353-8	Sequence 8, Appli
39	781	17.3	854	4 US-09-619-353-10	Sequence 10, Appli
40	756	16.7	669	4 US-09-361-631-7	Sequence 7, Appli
41	683.5	15.1	1058	2 US-08-687-289A-5	Sequence 5, Appli
42	672.5	14.9	872	3 US-08-337-797A-2	Sequence 2, Appli
43	672.5	14.9	872	3 US-08-258-523-2	Sequence 2, Appli
44	622	13.7	912	4 US-08-617-785-2	Sequence 2, Appli
45	622	13.7	912	4 US-09-641-318-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-09-134-513-2
Sequence 2, Application US/09134513
Patent No. 6210964
GENERAL INFORMATION:
APPLICANT: Brown, Edward M.
APPLICANT: Diaz, Ruben
APPLICANT: Bai, Mei
APPLICANT: Quinn, Stephen J.
TITLE OF INVENTION: The Avian Extracellular Calcium-Sensing
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Vinson & Elkins L.L.P.
STREET: 1455 Pennsylvania Avenue, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: U.S.
ZIP: 20004-1008
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/134,513
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sanzo, Michael A.
REGISTRATION NUMBER: 36,912
REFERENCE/DOCKET NUMBER: BRI331/13003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)639-6585
TELEFAX: (202)639-6604
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1059 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: protein
HYPOTHETICAL: NO
US-09-134-513-2

Query Match 24.1%; Score 1089.5; DB 4; Length 1059;

Best Local Similarity 34.4%; Pred. No. 1.7e-95;
Matches 266; Conservative 157; Mismatches 350; Indels 103; Gaps 23;

26 SQLRMKGDDYVLGGLPLG---EAEAGLRSTRPSSPVCTRFSSNGLLMALMKMAYEE 82

Db	23	NORAKOKEDILIGGIFRHHFGVAAKODOLKS - PRESVECIARYNFGFRMLQAMIFAEE	80
Qy	83	INNSDILPGRLCYDLEFDTCSEVYVAMKPSLMLA -- KAGSRDIAAYCANTQYQAPYLA	140
Db	81	INNSPNLLPNNTLGLRFDTCNTYSKALEALISFVAONKIDLSLDFECNCEHPISTIA	140
Qy	141	VIGHPSSELAWYTKFSEFSEFMPOVSYGASMELLSARETEPSPFRTPYPSRVOLTADEL	200
Db	141	VVGATGSGSVTAIVANLGLFIYIPQVSYAASSRSLSNKNOFKSFLRTIPNDEHQATAMADI	200
Qy	201	LOEGEMNNVVALSGDDEYGRGSLFISMLAARICIAHEBYLPLPADSRSLKAVDYL	260
Db	201	IEYFMNNVGTIADDDGRGRIEKFREAEERICIDFSELI --- SQYSDDEEIOQV	260
Qy	261	HOVNOSSVOYVLLFASVAAHALNENYSISRSLSKVVAVASEAMTSDLVMGIPMAOMGT	320
Db	257	EVIQNSTARVIYVSPSGDLEPLKEIYRNRITKIMLASAMASSSLIAMPEFRYIGS	316
Qy	321	VLGFLQRAOQHERRPOY - KTHLALANDPAPCSALGE -----	365
Db	317	TIPLRLKAGQIPGFRREFLOKVKPKKSANNGAKEFEWETENCYLPJSEKSNPASASEPKA	376
Qy	357	REOGJ ----- EEDVVGQRPCDCITLONYSAGLNHHQTESVYAAYVSA	401
Db	377	HEEGJAGNGNTPAARPCPTGDENTISVETPMDFTHLR ----- ISYNIYAVYSIA	427
Qy	402	QALNHTLOCNA ----- SGCPAADPYKPMQOLENNMTNLTFPHNG - GLPLRFDSGANDME	453
Db	428	HALDDIYTCPTBKGLFTNGSCADIKRYKAEAMOVILAKHLNHLNTSNMGBOYVDFDEFGDLYGN	487
Qy	454	YDLKLM -- WVOGSVPLRHDYGRFN - SLIPLRERL ----- KIRWHTSDNOKPVYSCSRQC	504
Db	488	YSIIMHLSPEDGSV - FEEYGHNYAKKERELFIENKTKIMSGFSKEVPEFNSCSDCL	546
Qy	505	EGOYVR - VKGHSCCYDCVDCBAGSYRONPDDIACYCGODEMSPERSTKCFRRRSFELA	563
Db	547	PGRTRGJIEGPEPTCCFEBCVCDPDEYSDETPASCDCDCPDEYMNENHTSCIPKOIEFLS	606
Qy	564	WGEPRVLLLLLLLSLALGLYLAALGFNHHSDPLVYOASGGPLA --- CFGYVGLGVCL -	619
Db	607	WTEPRGIMLTIFAYLIGITLTFVYLGVTTKFKNPTIYVATNRNLSYLLFSILC --- CFS	666
Qy	620	SVLLEPRQPSARCLAOOPLSHLPLTGLSTLFLQAAEIFY -- ESELPLSMADRLSGCLR	677
Db	663	SSLFETIGEPOMWTKRLRQAPAGISFVULCISCIILVKTNRVLLVEFAKIPTS ----- LH	714
Qy	678	GPW ----- AMLVVLLAMILEVALCTWTVLVAPEPVYVDMHLLPTEAL - VHCRTRSWVSFG	731
Db	715	RKMGMNLQFLVYLCFTYQVYVICVIMLYTAPESPSTYR - HELEDEILFTICHESMALG	773
Qy	732	LAHATNATLAFCLFGLTFVLYRSQPCYNARAGLTPEAMLAFTIYWSFVPLLANOYVLLP	791
Db	774	FLIGYTCILMAICEFFAFKSKRLEPENENAKETIIFSMILIEFIWISIPAYASVYGFYS	833
Qy	792	AVOMGALLICVILITLAAPHLRQVCLLRKOGCLNMP	827
Db	834	AVEVATILTAASFGLLACIFENKVYIILFKFSRNTIE	869

RESULT 2
 US-08-485-568-7
 Sequence 7, Application US/08485588
 Patent No. 5688938
 GENERAL INFORMATION:
 APPLICANT: Edward M. Brown
 APPLICANT: Steven C. Hedert
 APPLICANT: Forrest H. Fuller
 APPLICANT: James E. Garrett, Jr.
 TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 TITLE OF INVENTION: MOLECULES
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:

```

ADDRESS: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US-08/485_588
FILING DATE: 7 June, 1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA: Including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991

ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/005
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEX: (213) 955-0440
TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1078 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-588-7

Query Match      23.6%, Score 1068.5; DB 1; Length 1078;
Best Local Similarity 30.3%; Pred. No. 1.8e+93;
Matches 269; Conservative 156; Mismatches 366; Indels 97; Gaps 26;

QY   12 WALL-----HPGTGAPLCSQQRLRMKGDVYLGLGFPLG---EAEELGLASRTPRSSPVCTR    64
      |||         | :|||          | :|||          | :|||          | :|||
Db   8 WLLALHTWTSAYGP---DORAQRKKDDITILGLPFIHFGVAAKDDDLKS--RPSEVEICIR    62
      ::|||       ||::|||       |||:|||       |||:|||       |:|:|:|:|~|

QY   65 FSSNCLMALMKMAVEIKNKSDLLPELRGRGYDFDTCCSERVVAMKRSMFLA--KACS    122
      ::|||       ||::|||       |||:|||       |||:|||       |:|:|:|:|~|
Db   63 YNFGEFRWLQMIRAIEIENISSPALPNLTUGRYLFDCNTVSKALETLSFAONKIDS    122
      ::|||       ||::|||       |||:|||       |||:|||       |:|:|:|:|~|

QY   123 RDIAACMYTQOPRVLAIVGVPHSSELAMVTGKFSPFILMPDVSYGASMELLISARETPPS    182
      :||:|:|:|~|:|:|:|:|~|:|:|~|:|:|~|:|:|~|:|:|~|:|:|~|:|:|~|:|:|~|
Db   123 LINDFEFCNCSEHIPSTIANVGATSGYSTAVANILLGLEFYTIPOVYSASSRSLSNKNQPKS    182
      ||:|:|:|~|:|:|:|~|:|:|~|:|:|~|:|:|~|:|:|~|:|:|~|:|:|~|:|:|~|

QY   183 FRRVRPSDRVOLTAIAELLDEEGGNWVALGSDDEYGKGQLSIFSALLAAAGICTAHBGL    242
      ||:|:|:|~|:|:|:|~|:|:|~|:|:|~|:|:|~|:|:|~|:|:|~|:|:|~|:|:|~|
Db   183 FLRIIPNDENHQTAMADIETEFRRMWVGTLAADDDYGRPGLEKEEFELEEEDRIDCFIDFSEL    242
      ||:|:|:|~|:|:|:|~|:|:|~|:|:|~|:|:|~|:|:|~|:|:|~|:|:|~|:|:|~|
```



```

OY 243 VPLPRADDSRLGKQVODVLHQVNOSSVOYVLLFASVAAHALFNYSISSRLSPKVVASAA 302
DB 243 I-----SQTDEBEIOHVVAVIONSTAKVIVVSSGPDLEPLIKEIVRRNITKIMLASBA 298
OY 303 WLTSIDLVMKLPMAOMGVTLGFLQGAOLHEPQYV-KTHLALADPAFCALSAGERE--- 358
DB 299 WASSSLIAMPQYFHVYVGGTIGFALKAGQIPGRREFLKVKHPKSVHNGAFKFEWETENC 358
OY 359 -----OGIEEDVVGORCPQCD-----CITLQNSA-----GLNHQ- 389
DB 359 HLOEBAKGLPVDITFLRGHEES--GDRFNSNSTARPLCTGDNENISSVETPYIDTTHLRI 416
OY 390 TFSVAAVSVAAQALHNTLQ-----NASGPAODPVKPMOLLENNYNTLF-HVGL 440
DB 417 STNVILAVYSIAHALODITYCLPGKGLFNGS-CADIKKVEAMQVLKHLRHLNFTNNNGE 475
OY 441 PLRFDSSGNDMEYDLKLM--VMQGSVPLRHDVGRFN-GSLRTERL-----KIRWHTSD 491
DB 476 QVTFDECGDLVGNYSILIMWHLSPEDGSIV-FKEVGYNNYAKKGERLFTINEKILMSGFS 534
OY 492 NOKPVSRCRQCEQOVR-VKGFHSCCYDCVCEAGSTRONPDIACTFCGODEWSPER 550
DB 535 REVPSNCSRDIACTRGKILGEPTCCFCEVECPDEGYSDETDASACNKCDDPMWSEN 594
OY 551 STRCFRRSRFLAMEPAVILLLLSTLALGLVLAALGLFVHHRDSPICVQASGFLA--- 607
DB 595 HTSCIAKELEFLSTWEPGIALTLPAVLGIFLTAFLVGLFIFKRNPIYKATNRRLSTYL 654
OY 608 CGLVCLGLVCL-VYLPPGQSPARCIAOPLSHPLTGLSTLFLQAEIFV--ESEL 664
DB 655 LFLSLC-----CRSSLEFLGEFQDWTCLRQPAFGISFLVCLISLIVKNRVLVFEAKI 710
OY 665 PLS-----WADRLSCILNGPMANLYVLLAMVEVALCTWYLVAFPEVYTDHMLPTEL 719
DB 711 PTFEHRKWMGLNQ-----FLVFLCTEMQIVICVIMLTAPSSYRNOLDEILF 762
OY 720 VHCRTSRWVSFGLAHATNATLAFCLGLFLVRSOPGCYNRARGTLFAMLAFITWSPV 779
DB 763 ITCHGSLMALAGFLGYCYCLAALCFPAFKSRKLPENFNEKFTFSILIFITWISFI 822
OY 760 PLANVOYVLRPAVOMGALLICVLGILAAFLHPLRCYLLMRQGLNTP 827
DB 823 PAVASTYKFSVAVEVIAIILAAISGLACIFENKITIILFKPSRWTE 870

RESULT 3
US-08-484-565-7
; Sequence 7, Application US/08484565
; Patent No. 5763569
; GENERAL INFORMATION:
; APPLICANT: Edward M. Brown
; APPLICANT: Steven C. Hebert
; APPLICANT: James E. Garrett, Jr.
; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
; MOLECULES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: First Interstate World Center
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FASTSEQ
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,565

```

```

; FILING DATE: 7 June, 1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below: 9
; APPLICATION NUMBER: 08/353,784
; FILING DATE: 9 December, 1994
; APPLICATION NUMBER: PCT/US/94/12117
; FILING DATE: 21 October, 1994
; APPLICATION NUMBER: U.S. 08/292,827
; FILING DATE: 23 August, 1994
; APPLICATION NUMBER: U.S. 08/141,248
; FILING DATE: 22 October, 1993
; APPLICATION NUMBER: U.S. 08/009,389
; FILING DATE: 23 February, 1993
; APPLICATION NUMBER: U.S. 08/017,127
; FILING DATE: 12 February, 1993
; APPLICATION NUMBER: U.S. 07/934,161
; APPLICATION NUMBER: U.S. 07/834,044
; FILING DATE: 11 February, 1992
; APPLICATION NUMBER: U.S. 07/749,451
; FILING DATE: 23 August, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Heber, Sheldon O.
; REGISTRATION NUMBER: 38,179
; REFERENCE/DOCKET NUMBER: 213/006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1078 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-484-565-7

Query Match 23.6%; Score 1068.5; DB 1; Length 1078;
Best Local Similarity 30.3%; Pred. No. 18e-93;
Matches 269; Conservative 156; Mismatches 366; Indels 97; Gaps 26;

OY 12 WALL-----HGTGAPICLSQOLRMKGDYVGLGFLPG---EAEAGLRRTSPSPVCR 64
DB 8 WYLLALTWHTSAVGP---DQRAQKGDITLGLFLPHFGVAAKQDQDKS--RPESECR 62
OY 65 FSSNGILMALAMKMAVEEINNKSDLLPGLRLGYDLFTDCESEPVVAMKPSIMFLA--KAGS 122
DB 63 YNRFGRFWLQAMIFALEEINSSPALLPNTLGLRIFDTCTNTVSKALEATLSFAQNKIDS 122
OY 123 RQIAVYCNTOYQPRYLANIYGHSSSLAMVTKFSPFLMPQVSGASSELARSPTFS 182
DB 123 INLDEFCNCEHPISTIAVAGATGSGVSTAVANLGLFYIPQVSYASSSRLISNNQFES 182
OY 183 FFRTPSDRVQTLTAAELLQEEFGMMVVALGSDDEYRGOLSFSAIARIGICIAHEGL 242
DB 183 FLRTIPNDHQATAMADILEYFRMMVGTIAADDYGRGCIKFRNEAERIDICIDFSL 242
OY 243 VPLPRADDSRLGKQVODVLHQVNOSSVOYVLLFASVAAHALFNYSISSRLSPKVVASAA 302
DB 243 I-----SQTDEBEIOHVVAVIONSTAKVIVVSSGPDLEPLIKEIVRRNITKIMLASBA 298
OY 303 WLTSIDLVMKLPMAOMGVTLGFLQGAOLHEPQYV-KTHLALADPAFCALSAGERE--- 358
DB 299 WASSSLIAMPQYFHVYVGGTIGFALKAGQIPGRREFLKVKHPKSVHNGAFKFEWETENC 358
OY 359 -----OGIEEDVVGORCPQCD-----CITLQNSA-----GLNHQ- 389
DB 359 HLOEBAKGLPVDITFLRGHEES--GDRFNSNSTARPLCTGDNENISSVETPYIDTTHLRI 416
OY 390 TFSVAAVSVAAQALHNTLQ-----NASGPAODPVKPMOLLENNYNTLF-HVGL 440
DB 417 STNVILAVYSIAHALODITYCLPGKGLFNGS-CADIKKVEAMQVLKHLRHLNFTNNNGE 475

```

```

Db 417 SYNYTIAVYSIAHLODIITCLPGRGLFTNGS-CADIKKEAVOMVLAHLNLTNNKGE 475
QY 441 PLRFDSSGNVMEYDLKLM--VWQGSVPLHDVGRFN-GSLRTERL-----KIRWHTSD 491
QY 476 QVTFDECDGLVGNYSIINMHLSPEDGSIV-EKEGYYNVAKKGERLFINDEKILMSGFS 534
QY 492 NQRPVSRCSRCQCGQVRR-VKGFHSCCYDCVCEASYSQNPDDIACFCGODEMSEPER 550
Db 535 REVPFNSCRDCLAGTRGILIEGPTCCFECVCECPDEYSEIETDASACNKCPCDFWSENEN 594
QY 551 STCRFRRSRLRAMEGEPAVILLTLLLSLALGLVLAALGLFVHHSDPVSAGGPIA--- 607
Db 595 HTSCIAKEIEFLSWTERPGIALLTFAYLGITLTFVGLGFTKFRTPYVAKNRELSTLL 654
QY 608 CEGIVLCGLVCL-SVLLFPQGPSPARCLAOPLSHLPYGLSLTFLQAAEIV--ESEL 664
Db 655 LFSILC-----CFSSSLFPIGEPQDMTCRLQPAFGISFVLCISCLIVTNVLLVFEAKI 710
QY 665 PLS-----WADRLSGCLRGPMAMLYVLLAMVEVALCTWYLVAPPEVYDWMHLPTEAL 719
Db 711 PTFHRRKMWGINTQ-----FLVFLCTFMQIVICIMLYTAPSSSYRQOELEDEITF 762
QY 720 VHCRTSRWVSFGLAHATUNATLAFICFLGTFLVRSOPGCGYNRARGLTFAMLAFITWVSFV 779
Db 763 ITCHEGSLMAGELIGYTCLLAICFPFAFKSRKLPERENAKFTTSMLEFFIWIWISFI 822
QY 780 PLANVOVILPRAVOMGALLCVLGLLAFLHLPYCYLLMRQPGINTPE 827
Db 823 PAVASTGKFEVSAVEVIALIILASFGILACIFNKIYIILFKPSNHTIE 870

```

RESULT 4

US-08-480-751-7
Sequence 7, Application US/08480751

Patent No. 5858684
GENERAL INFORMATION:

APPLICANT: Edward F. Nemeth
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: Forrest H. Fuller
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
TITLE OF INVENTION: MOLECULES
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071

COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/480,751
FILING DATE: 7 June, 1995
CLASSIFICATION: 435

PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993

```

APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Hebert, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 213/004
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1078 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-480-751-7

```

Query Match 23.6%; Score 1068.5; DB 2; Length 1078;
Best Local Similarity 30.3%; Pred. No. 1.8e-93;
Matches 269; Conservative 156; Mismatches 366; Indels 97; Gaps 26;

```

QY 12 WALL-----HPTGAPLCLISQOLRRKGDYVIGLPLG----EAEPAIGRSRRSPVCTR 64
Db 8 WVLLALTWHTSAYGP---DQRAOKRGDIIIGLGLPIHGVAAKQDDIKS--RPSVACIR 62
QY 65 FSSNGILMALAMKAAVEINNKSDLLGLRLGYLPTDCSEPVYAMPSIMFLA--RAGS 122
Db 63 YNFGFRLQAMTARIEINSSPALRLNLTGYITFCNTVSALEATISFAVQNKIDS 122
QY 123 RDIAYCNTQYOPRVLAIVGPHSSSELMWTKFSEFLMPQVSYGASMEILSARETFPS 182
Db 123 LINDFCNCSHIPSTIAVAVGATGSGVSTAVANLGLFYIIPQVSYASSRILSKNQFS 182
QY 183 FFRIVPSDRVQTLAAELDFEGNNWYAAAGSDDEYGRGICSTISALAAAGCITIAEGL 242
Db 183 FLRTIPNDEQATMAIDIEFRFNNWGTIAADDYGRPGIEKREBAEERDIDISEL 242
QY 243 VPLRADSRIGKYODVILHOVNOSSVOVILLFASVHAHAHLFNSSISRLSPKYVASEA 302
Db 243 I-----SQYSDDEEIQHYVEVYQNSTAVIYVSSGPDLEPLIKETIVRNITGKIWLASEA 298
QY 303 WLTSIDLVMGLPGMAQMGTVLGLQGAQLHEFPQYV-KTHLALATDPAFCASALGERE-- 358
Db 299 WASSSLIAMPQYFHVVGSTIGFALKAGQIPQFRFLKKBVHGRKSVHNGFAKEFEETFCNC 358
QY 359 -----QGLEEDVYGQRCPQD-----CITLQVNSA-----GLNHHQ- 389
Db 359 HLOGAKPLPVDTFLGLGHBS--GDRFNSNSTAFRLCAGDENISSEVEPYIDYTHIRI 416
QY 390 TFSYAAVYSVAQALHNTLQC-----NASGCPADDPKPMQMLNMYNLFP--HYGGL 440
Db 417 SYNYTIAVYSIAHLODIITCLPGRGLFTNGS-CADIKKEAVOMVLAHLNLTNNKGE 475
QY 441 PLRFDSSGNVMEYDLKLM--VWQGSVPLHDVGRFN-GSLRTERL-----KIRWHTSD 491
Db 476 QVTFDECDGLVGNYSIINMHLSPEDGSIV-EKEGYYNVAKKGERLFINDEKILMSGFS 534
QY 492 NQRPVSRCSRCQCGQVRR-VKGFHSCCYDCVCEASYSQNPDDIACFCGODEMSEPER 550
Db 535 REVPFNSCRDCLAGTRGILIEGPTCCFECVCECPDEYSEIETDASACNKCPCDFWSENEN 594
QY 551 STCRFRRSRLRAMEGEPAVILLTLLLSLALGLVLAALGLFVHHSDPVSAGGPIA--- 607
Db 595 HTSCIAKEIEFLSWTERPGIALLTFAYLGITLTFVGLGFTKFRTPYVAKNRELSTLL 654

```

QY 608 CGLVCLGLVCL-SVLEFPQSPARCLAQPLSHLPLTGLSTFLQAAELFV--ESEL 664
 Db 655 LFSLLC-----CFSSSLFETGEPODWTCKLROPARGISFVLICISCLVNTKRNVLVFEAKI 710
 QY 665 PLS-----WADRLSGCLGPNAMLVLLAMVEVALCTWYLVAPPEVVTDMHMLPTAL 719
 Db 711 PTFRRKMWGINTLQ-----FLVFLCTFMQIVICIMYLTAPSSRYRNOLEDELIIF 762
 QY 720 VHCRTSRVSGFLAHATNATLAFCTGTLVRSOPGCGYNRARGITFAMLAFTWVSFV 779
 Db 763 ITCHEGSLMAGFLIGYTCCLAICFFPAFKSRKLPENFNKAKFTFEMLIFFIWMISFI 822
 QY 780 PLANVOYVLRPAVOMGALLICVGLIAAFLHPLRCYILMROPGINTPE 827
 Db 823 PAYASTYKRFVSAVEVYIAILASFGLLACIFENKIYIILFKPSRNTIE 870
 RESULT 5
 US-08-943-986-7
 : Sequence 7, Application US/08943986
 : Patent No. 5962314
 : GENERAL INFORMATION:
 : APPLICANT: Edward M. Brown
 : APPLICANT: Steven C. Hebert
 : APPLICANT: James E. Garrett, Jr.
 : TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 : TITLE OF INVENTION: MOLECULES
 : NUMBER OF SEQUENCES: 20
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Lyon & Lyon
 : STREET: First Interstate World Center
 : STREET: Suite 4700
 : STREET: 633 West Fifth Street
 : CITY: Los Angeles
 : STATE: California
 : COUNTRY: USA
 : ZIP: 90071
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
 : COMPUTER: IBM PC compatible
 : OPERATING SYSTEM: PC-DOS/MS-DOS
 : SOFTWARE: FASTSEQ
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/943,986
 : FILING DATE: 03-OCT-1997
 : CLASSIFICATION: 530
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/484,565
 : FILING DATE: 7-June-1995
 : APPLICATION NUMBER: 08/353,784
 : FILING DATE: 9 December, 1994
 : APPLICATION NUMBER: PCT/US/94/12117
 : FILING DATE: 21 October, 1994
 : APPLICATION NUMBER: U.S. 08/292,827
 : FILING DATE: 23 August, 1994
 : APPLICATION NUMBER: U.S. 08/141,248
 : FILING DATE: 22 October, 1993
 : APPLICATION NUMBER: U.S. 08/009,389
 : FILING DATE: 23 February, 1993
 : APPLICATION NUMBER: U.S. 08/017,127
 : FILING DATE: 12 February, 1993
 : APPLICATION NUMBER: U.S. 07/934,161
 : FILING DATE: 21 August, 1992
 : APPLICATION NUMBER: U.S. 07/834,044
 : FILING DATE: 11 February, 1992
 : APPLICATION NUMBER: U.S. 07/749,451
 : FILING DATE: 23 August, 1991
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Hebert, Sheldon O.
 : REGISTRATION NUMBER: 38,179
 : REFERENCE/DOCKET NUMBER: 213/006
 : TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 7:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1078 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-943-986-7
 Query Match 23.6%; Score 1068.5; DB 2; Length 1078;
 Best Local Similarity 30.3%; Pred. No. 1.8e-93;
 Matches 269; Conservative 156; Mismatches 366; Indels 97; Gaps 26;
 QY 12 WALL-----HPTGAPRLCSQQLRMKGDVYLGLEPLG--EAENAGLSRRPSPVCTR 64
 Db 8 WVLATLWHTSAYGP---DRAOKKGDIIILGLPIHGVAAKQDDLS--RPSVEICIR 62
 QY 65 FSSNGLMALAMKMAVEIRNKSDLLPGLRLGYDLDFOTCSEPVYAMKSLMFLA--KAGS 122
 Db 63 YNFKGFRLQAMTFAIEIRINSSPALPLMLTGLYRIEDCNVSKALEYTSFVQNKIDS 122
 QY 123 RDIAYCNVTOYOPRVLAIVIGPHSSELMVTKFEFSFIMPQVSYGASMEILSARETPS 182
 Db 123 LNLDEFNCSEHISTIAVVGATGCVSTAVANLLGLFIYIQVSYASSRLSNKQPKS 182
 QY 183 FFRVPSDROVLTAAELLQEPGMNVAALGSDDEYGRGSLISALANGICIAHEGL 242
 Db 183 FLRTIPNDHQATAMADIEFEYFMNWGTIAADDYGRPGIEKFEFEAEERDIDCESEL 242
 QY 243 VPLPADSRILGKQODVHOVNOSVOYVLLFASVHAHALFNSSIRLSRPKWVNSEA 302
 Db 243 I-----SOYSDEEIOHVVEVIONSTAKIVVSSGPDLEPLKEVYRNITGKMLASEA 298
 QY 303 WLTSDLVGLPGMOMGTVLGFLQGAOLHEFPQYV--KTHLALATDPAFCSALGERE--- 358
 Db 299 WASSSLIAMPQYFHYVGGTIGFALKAGQIPGRFELKVVHPRKSVHNGFAEFMEYETNC 358
 QY 359 -----QGLEDDYVQRCPOCD-----CITLQVSA-----GLNHQ- 389
 Db 359 HLOEGAKGPLPVDFTLRGHEBS--GDRFSNSSTAFRLPCLTDENLSVEYPIDYTHIRI 416
 QY 390 TFSYAAVYSVAOALHNTLOC-----NASGCAQOPVAPWOLLEMYNLTFF--HVGL 440
 Db 417 SYNTYLAIVSTAHADLIYTLPGRGLETNGS--CADIKVAMQVYKLRHLNFTNNNGE 475
 QY 441 PLRFDSSGNVDMEXDLKIM---VMQGSYRLHDVGRFN--GSLRTERL-----KIRWHTSD 491
 Db 476 QVTFDECGDLVGNYSIIMWHLSPEDGSIV--FKVEGYVNVYAKKGRFLINBEKILMSGFS 534
 QY 492 NQKPVSRCSROCGQVNR--VKGRHSCCYDVCDEAGSYRONPDIACITPCGODEMSER 550
 Db 535 REVPFNSCRDLGTRGIIIEGEPTCCFCEVCEPDGEYSDETASACNKPDPDEFWSEN 594
 QY 551 STFRFRSRFLANGEPVALLLLLSLALGLVLAALAEFLVHNHRSAPLVQASGGLA--- 607
 Db 595 HTSCAKIEIFLSTWEPGIALITFAVIGITLFAVIGVIFKFRNTPYVKAFTNELSL 654
 QY 608 CGLVCLGLVCL-SVLEFPQSPARCLAQPLSHLPLTGLSTFLQAAELFV--ESEL 664
 Db 655 LFSLLC-----CFSSSLFETGEPODWTCKLROPARGISFVLICISCLVNTKRNVLVFEAKI 710
 QY 665 PLS-----WADRLSGCLGPNAMLVLLAMVEVALCTWYLVAPPEVVTDMHMLPTAL 719
 Db 711 PTFRRKMWGINTLQ-----FLVFLCTFMQIVICIMYLTAPSSRYRNOLEDELIIF 762
 QY 720 VHCRTSRVSGFLAHATNATLAFCTGTLVRSOPGCGYNRARGITFAMLAFTWVSFV 779
 Db 763 ITCHEGSLMAGFLIGYTCCLAICFFPAFKSRKLPENFNKAKFTFEMLIFFIWMISFI 822
 QY 780 PLANVOYVLRPAVOMGALLICVGLIAAFLHPLRCYILMROPGINTPE 827

Db 823 PAVASTGKFSAYEVTAIIAASFGLLACJFFENKIYIILKPSRNTIE 870

RESULT 6

US-08-353-784-7

Sequence 7, Application US/08353784

Patent No. 6011068

GENERAL INFORMATION:

APPLICANT: Edward F. Nemeth, Edward M.

APPLICANT: Brown, Steven C. Hebert,

APPLICANT: Bradford C. Van Wageningen, Manuel

APPLICANT: F. Balandrin, Forrest H. Fuller,

APPLICANT: Eric G. Delmar, and Scott T. Moe

TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: First Interstate World Center

STREET: Suite 4700

STREET: 633 West Fifth Street

CITY: Los Angeles

STATE: California

COUNTRY: USA

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: FASTSEQ

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/353,784

FILING DATE: 9 December, 1994

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

PRIOR APPLICATION DATA: including application

PRIOR APPLICATION DATA: described below: 8

APPLICATION NUMBER: PCT/US/94/12117

FILING DATE: 21 October, 1994

APPLICATION NUMBER: U.S. 08/292,827

FILING DATE: 23 August, 1994

APPLICATION NUMBER: U.S. 08/141,248

FILING DATE: 22 October, 1993

APPLICATION NUMBER: U.S. 08/009,389

FILING DATE: 23 February, 1993

APPLICATION NUMBER: U.S. 08/017,127

FILING DATE: 12 February, 1993

APPLICATION NUMBER: U.S. 07/934,161

FILING DATE: 21 August, 1992

APPLICATION NUMBER: U.S. 07/834,044

FILING DATE: 11 February, 1992

APPLICATION NUMBER: U.S. 07/749,451

FILING DATE: 23 August, 1991

ATTORNEY/AGENT INFORMATION:

NAME: Hebert, Sheldon O.

REGISTRATION NUMBER: 38,179

REFERENCE/DOCKET NUMBER: 209/069

TELECOMMUNICATION INFORMATION:

TELEPHONE: (213) 489-1600

TELEFAX: (213) 955-0440

TELEX: 67-3510

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 1078 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-353-784-7

Query Match 23.6%; Score 1068.5; DB 3; Length 1078;

Best Local Similarity 30.3%; Pred. No. 1,8e-93;

Matches 265; Conservative 156; Mismatches 366; Indels 97; Gaps 26;

QY 12 WALL-----HPGTAPLCLSQLRNNKGDYVLAGLEPLG---EAEFAGLRSTRPSSPCTR 64

Db 8 WVLLALTWHTSAYGP---DQRAQKGDIIILGGLPIHFHGAQAQDQDKS--REPSEVICIR 62

QY 65 FSSNGLLMALAMKNAVEINKKSDLLPGIRGYPDPTCSPPVAMPSIMFLA--KAGS 122

Db 63 YNERGFRLQAMTFALIEINSSPALNPLTLGYIFDTCNTVSKALNATLSFAQNKIDS 122

QY 123 RDIAYCNTQYOPRVLAIVGPHSSSELAHWYGFSEFLMPOVSYGASMEIARSRETPS 182

Db 123 LNDEFPCNSHIPSTIAVVGATGSGVSTAVANLGLFYIPQVSYASSRLSKNFKS 182

QY 183 FETVPDSRVQVLAIAELLOFGNNWYAALGSDDEYRQGLSIFSAALAGICIAHEGL 242

Db 183 FLRTIPNDEHQATAMADIIIEYFRNNWYGTIAADDYGRPGIEKREAEERDIDISEL 242

QY 243 VPLPRADSRGKQVDVLAHQVNOSSQVVLVAFVHAHNLFNYSISRLSPKVVVSEA 302

Db 243 I---SOYSDDEETOHVEVYQNSTAVIVFSSGPDLEPLIKETIRRNITGKIWLASEA 298

QY 303 WLTSDLVWGLPGMAQMGTVLFGQALHEFPQYV-KTHLATATDAPCSALGERE-- 358

Db 299 WASSSLIAMQYFHYVGGTIFGALKAGQIIGREFLAKVHRKSVHNGFAKEFWDETFNC 358

QY 359 -----OGLEEDVVGRCPOCD-----CITLQNVSA-----GLNHQ 389

Db 359 HLAGKAGPLFVDPTFLRGHEES--GDRFNSSTAFRLPGTCDEWISSVETPYIDYHLRI 416

QY 390 TFSYTAAYVSYAOLNHTLQ-----NAGSPADDPKPMQMLNMTNLT-FHNGL 440

Db 417 SYNVYLAIVYSTAHALODIYCLPGRGLFTNGS-CADIKKVAQVLAHLRLNNTNNGE 475

QY 441 PLRFSSGNVMEYDLKLM--VWQGSVPRLDHGRN-GLRFRERT-----KIRWHTSD 491

Db 476 QVTFDECGDLVGNKSIINMHLSPEDGSIV-FKEGYINNYAKKGERFLFINEKILMNGFS 534

QY 492 NQRFVSKSRQCGQYVR-VKGFHSCCYDVCDEAGSYRONPDDIACCTGODEMSPER 550

Db 535 REVFVSNCSRCLAGTRKGIIEGEPCCFECECPDEYSDETDASACNKCPCPDFGNGEN 594

QY 551 STFRFRSRRLAMGEPAVILLILLLSALGLVLAALGLFVHNDSPLYOASGSPLA--- 607

Db 595 HTSCIAKEIEFLSWTEPPGIALTLFANVLGIFLAFVGLVFKFNNTIYVATNRELSTYL 654

QY 608 CFGVLVCLGLVCL-SVLLFPGQPSPARCLADQPLSHLPGLTGYLFLQAAEIFV--ESEL 664

Db 655 LFSLLC-----CFSSSLFIFIGEPQDMTCRLRQPARGISFVLCISCIIVKTRVILLVFEAKI 710

QY 665 PLS-----WADRLSGCLRGPMAMLVYLLAMLEVVALCTWYLVAFPEVYTDHMLPTREAL 719

Db 711 PTFSHRRKMGINTQ-----FLVFLCJFMQIVICVIMLYTAPPSYRNOLEDEIIF 762

QY 720 VHCRTSMVSHGLAHATNATLAFICFLGTFLVRSQPCVYNARAGLTAMLAFTWVSFV 779

Db 763 ITCHEGSLMALGLIGITCLLAICFFPARSKRLPENFEAKITISMLFTWLSFI 822

QY 780 PILANVOVLRPAVOMGALLVGLIILAFLTPRCYLLMRPGIANTPE 827

Db 823 PAVASTGKFSAYEVTAIIAASFGLLACJFFENKIYIILKPSRNTIE 870

RESULT 7

US-08-484-7198-7

Sequence 7, Application US/084847198

Patent No. 6031003

GENERAL INFORMATION:

APPLICANT: Edward F. Nemeth, Edward M.

APPLICANT: Brown, Steven C. Hebert,

APPLICANT: Bradford C. Van Wageningen,

APPLICANT: Manuel F. Balandrin,

APPLICANT: Forrest H. Fuller, Eric G.

TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE

```

: TITLE OF INVENTION: MOLECULES
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon
: STREET: First Interstate World Center
: STREET: Suite 4700
: STREET: 633 West Fifth Street
: CITY: Los Angeles
: STATE: California
: COUNTRY: USA
: ZIP: 90071
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: MS Word
: SOFTWARE: FASTSEQ for Windows Version 3.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/484,719B
: FILING DATE: 7 June, 1995
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/353,784
: FILING DATE: 9 December, 1994
: APPLICATION NUMBER: PCT/US/94/12117
: FILING DATE: 21 October, 1994
: APPLICATION NUMBER: U.S. 08/292,827
: FILING DATE: 23 August, 1994
: APPLICATION NUMBER: U.S. 08/141,248
: FILING DATE: 22 October, 1993
: APPLICATION NUMBER: U.S. 08/009,389
: FILING DATE: 23 February, 1993
: APPLICATION NUMBER: U.S. 08/017,127
: FILING DATE: 12 February, 1993
: APPLICATION NUMBER: U.S. 07/934,161
: FILING DATE: 21 August, 1992
: APPLICATION NUMBER: U.S. 07/834,044
: FILING DATE: 11 February, 1992
: APPLICATION NUMBER: U.S. 07/749,451
: FILING DATE: 23 August, 1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Douglas C. Murdoch
: REGISTRATION NUMBER: 37,549
: REFERENCE/DOCKET NUMBER: 213/007
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (213) 489-1600
: TELEFAX: (213) 955-0440
: TELEX: 67-3510
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1078 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-08-484-719B-7
:
: Query Match 23.6%; Score 1068.5; DB 3; Length 1078;
: Best Local Similarity 30.3%; Pred. No. 1.8e-93;
: Matches 269; Conservative 156; Mismatches 366; Indels 97; Gaps 26;

```

```

: Db 183 FLRTIPNDHEQATAMADITTEFRMNVGTIAADDDYGRPGIKKEFEAEEDNDICDSEEL 242
: 243 VPLPRADSRRLGKQVDVHLQVNSSQVYVLLFASVAAHALFNYSISSRLSPKVVASEA 302
: 243 I----SQSDEEIQHVVEVIONSKAVIVFSSPDLEPLIKELVRNNTGKILASEA 298
: 303 WITSDLVNGLPGMAQMGTVLGLQAGQJHEFPQIV KTHALATDPAPCSALGERE--- 358
: 299 WASSSLIAMPQYFHVYGGTIGFALRAGQIPGRREPLKRVHBRKSVHNGFAKEFWETPNC 358
: 359 -----OGLEEDVVGRCPOCD-----CITLQNVSA-----GLNHHO 389
: 359 HLOEBAKPLPYDTFLRKHES--GDRFSNSSTARPLCTGDNENISSVETPIIDTILRI 416
: 390 TFSYAAVYSVAQALHNTLQC-----NASGCAQDPYKRMOLLENNYILTF-HVGL 440
: 417 SYNVYLAIVYSTAHALQDIYTCIPGRGLFTNGS-CADIKKVEAMQYLKHLRHLNFTNNMG 475
: 441 PLRPDSSGNVMEYDLKLM---VMQSVPLRHVDREN-GSILRTRL-----KIMHTSD 491
: 476 QVTFDECGDLVGNYSIIIMWHLSPEDGSIV-FREVGYYVYAKKGRRLPINEKILMSGFS 534
: 492 NOKPYRSGROCOEQGVNR-VKGFHSCCYDCVCEAGSYRQNDIACTFGQODEMSPER 550
: 535 REVPPSNCSROCLACTRGKGIIEGPTCCFEVCECDGEYSDETSACNKCDDPDMSEN 594
: 551 STCRFRRSRLAWGEPVALLLLLSLALGLVLAALGFVYHNRDPLVQASGGFLA--- 607
: 595 HTSCIARKIEFLSWTEPFGIALTFPAVIGIFLTAIVLVFIKRNTPYVAKTNRELSTYL 654
: 608 CFGVLCLGLVCL-STYLEPPGQSPARCQAQPLSHLPLTGLCLSTFLQAAELFV--ESEL 664
: 655 LFLSLC---CFSSSLEFPIGEQDWTCLROPAFGISVLCISCLVKTNRVLVFEAKI 710
: 665 PLS-----MADRLSGCLGPMAMLVLLAMLEVEVALCTWYLAAPPEVVTDMHMLPTAL 719
: 711 PTFPRKRWGUNIQ-----FLVFLCTFMQIYICVIMLYTAPRSSYRQDELEDELIF 762
: 720 VHCRTRSVNSFGLAHATNATLAFLCFLCTFLVRSOPGCIYNRAIGITPAMLAIVTWSFV 779
: 763 ITCHEGSIMALGFLIGYCTLAALTCFFPAFRSKRKLPEFNENKFTTFESMLTFEIVWISFI 822
: 760 PLLANQVVLPRAYOMGALLCVGIIAFLHPRCYILMRPGULTPE 827
: 823 PAVASTYKEVSAVEVIAIILASFGLLACIFFPNKIYITLFRPSNTIE 870
:
: RESULT 8
: US-08-484-159-7
: Sequence 7, Application US/08484159
: Patent No. 613146
: GENERAL INFORMATION:
: APPLICANT: Bradford C. Van Wagonen
: APPLICANT: Manuel F. Balandrin
: APPLICANT: Eric G. Del Mar
: APPLICANT: Edward F. Nemeth
: TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
: MOLECULES
: NUMBER OF SEQUENCES: 20
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Lyon & Lyon
: STREET: First Interstate World Center
: STREET: Suite 4700
: STREET: 633 West Fifth Street
: CITY: Los Angeles
: STATE: California
: COUNTRY: USA
: ZIP: 90071
: COMPUTER READABLE FORM:
: MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS

```

```

SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,159
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Hebert, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET NUMBER: 214/101
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1078 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-159-7

Query Match      23.6%  Score 1068.5  DB 4:  Length 1078:
Best Local Similarity 30.3%  Pred. No. 1.8e-93;
Matches 269; Conservative 156; Mismatches 366; Indels 97; Gaps 26;
QY 12 WALI-----HPGTGAPLCLISQRLRMKGDYVLGGLPPLG---EAEAGLRSRTSPSPVCTR 64
DB 8 WVLALATWHTSAVGR---DQRAQKKGDIIILGLFPIHGVAAKKODLKS--RPSVEGICR 62
QY 65 FSSNGLLMALAKMAVEINNKSLDLPGLRIGYDLFTGCEPVAAMKPSLMFLA--KAGS 122
DB 63 YNFGFRLMLOMTFAIEINSSPALLPLNLGRIPIFTCNVSKALNTATLSFVAQNKIDS 122
QY 123 RDIAYACNTYOYORVLAIVGPHSSELAMVTKFFSFILMPOVSYGASMSLLSARETPPS 182
DB 123 LNDIFENCESEHPISTIAVVGATGSGVSTAVANILGLFIYIPQVSYASSRLSKNPKS 182
QY 183 FERTVPSDRVQLTAAAELELQFEGNNWVAALGSDDEYSGROGLSIFSALAAAGICIAHEGL 242
DB 183 FLRTIPDEHQATAMADIIIEFRNNWGTIAADDDYGRPGIEKREAEEDIDIDISEL 242
QY 243 VPLPRADSRKGVQDVLHVOYNOSSVOVLLFASVHAHALEFNYSISSRLSPKYWASSEA 302
DB 243 I-----SOYSDEEEOHVEVYIONSTAKVIVFSSGPDLEPLIKETIVRRNITGKIWLASEA 298
QY 303 WLTSDLMWGLPGMAQMGTVLGLORGAOLHEFPQYV-KTHALATDPAFCALGERE--- 358
DB 299 WASSSLIAMPOYFHVVGCTIGFALKAGQIDGFEFLKRVHRRKSVHNGFAKEFEETENC 358
QY 359 -----OGLIEDVVGORCPQCD-----CITLONVSA-----GLNHQ- 389
DB 359 HLDGAGKGPLPVDFTFLGHEBS--GDRFNSNSTAFRPLCTGDENTISSVEPTIDYTHIRI 416

```

```

QY 390 TFSYAAVYSVAQALHNTLOC-----NAGCPADDPVAKPMOLNMYNLTFF-HVYGL 440
DB 417 SYNYLAVYSTAHMLQDIYICLPGRGLETPNS-CADIKRYAMQVYLHRLNLTNTNMG 475
QY 441 PLRFDSSGNVMEYDLKIM---VMQSVPLRHVDYGRN-GLRTERL-----KIRWHTSD 491
DB 476 QVTEDECDGLVGNYSIINWHLSPEDGSIV-FKEGYNNVAKKGERLFINKEKILMSGFS 534
QY 492 NOKRVSGSRQCCGQVRR-VKGFHSCYDQVDEASYSRONPDICTGCGODEMSPER 550
DB 535 REVPFNSCSDCLAGTKRKITGEPTCCFECVCPDDBISDPTDASACNKCDDPDSNEN 594
QY 551 STCFRRRSRRLWAGEPAVLLILLLSIALGLVLAALGLFVHNRDSPVYASGGPLA--- 607
DB 595 HTSCIAIEIEFLSWTEPEFGIALTLFAVLGIFLFAVGVGFIKFNPTLVKATNELSYLL 654
QY 608 CFGVCLGLVCL-SVLPFROPSPARCLAQOPLSHLPLTGLSLFLQAAEITV--ESEL 664
DB 655 LFSLLC---CFSSSLFETGEPQDWTQRLQRPAGISFVLICISILVKNRVLVLFPAKI 710
QY 665 PLS-----MADRLSGCLRGPMAMLVLLAMLVEALCTWYLVAPPEVYTDHMLPTREAL 719
DB 711 PTFHRRKMWGLNQ-----FLVFLCTFMQIVICVIMLYTAPSPSYRQDELEIF 762
QY 720 VHCRTSMVSFGLAHATNATLAFICLGTFLVRSQPCYNNRARGLFTAMLAFTTWSEV 779
DB 763 ITCHEGSLMALGFLIGYTCLLAALCFEFAFRSKLPEFNFAKRTFPSMLFFIWMISFI 822
QY 780 PLANNQVYLRPAVOMGALLCVGILTAFLPRCICYLLMRQPGUNTE 827
DB 823 PAVASTYKREYSAVEVYAILLAASFGLACIFENKRYITILEKPSRNTLE 870

RESULT 9
US-08-485-588-5
Sequence 5, Application US/08485588
Patent No. 5688938
GENERAL INFORMATION:
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: Forrest H. Fuller
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/485,588
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994

```

APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179
REFERENCE/DOCKET INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1085 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-485-588-5

Query Match 23.4%; Score 1060; DB 1; Length 1085;
Best Local Similarity 30.0%; Pred. No. 1.2e-92;
Matches 269; Conservative 155; Mismatches 362; Indels 112; Gaps 26;

QY 6 VGLSLMALHPGTAPLCLSLQGLKGVYGLFPL--GEA-EEAGLRSTRSPSPVC 62
DB 10 LAFSTWCHSAVGP-----DRAQKKGIILGGLPIHFGVAVKDQDLKS--RPSVRC 61
QY 63 TRFSSNGLLMALMKAAVEINNKSDLLPGLRGLDLPDTCSEPPVAMKPSIMFLA--KA 120
DB 62 IRYNFGFRWLQAMFALBEINSSPALHNPMLGRIIPDTCMTVSALATLSFAVQNKI 121
QY 121 GSRDLAAYCNTQYQPRVAVIGPHSSSLAMVTKFFSFLMPOVSYGSMELLSARETF 180
DB 122 DSLINDEFNCSEHPISTIAVVGATGSGISTAVANLGLFYIPVSYASSRLLSKNQF 181
QY 181 PSFRTVPDSRYOLPAALAELOFGNNVVAALGSDDEYGRGISTIFSALAAAGCIANE 240
DB 182 KSFRLTIPDEHOATAMADIEYFRNMVGTIAADDYGRPGIEKFRERAEEDICIDPS 241
QY 241 GLVPLPRADDSRLGKQVDVLAHQVNOSSVOVLLFASVHAHALFNYSISSRLSPKWVAS 300
DB 242 ELI-----SQYSDERKIQVVEVIONSTAVIYVSSGPDLEPLIKEIYRNRITGRITLAS 297
QY 301 EAMLSIDLVMGLPGMAQMGTVLGLTQGAOLHEFPQYV-KTHIALATDPAFCASAL----- 354
DB 298 EAMASSSLTAMPEYFHVGGTIGFGLKAGQIPGFFREFLOKVPKRSVHNGFAKEWEETE 357
QY 355 -----GEREGL-----BEDVVGQCGPCQDCITLON 380
DB 358 NCHLDGAGAPLPVDITFLNGHEGGRKLSNSPTARPLCTGEBNLISSVETPTMDYTHLR- 416
QY 381 VSAGLNHGTFSVYAAYSVAQALNHTLOC-----NASGCPADDPYKPMQOLENNMTN 432
DB 417 -----ISYNYLVAVYSIAHALDITVICIPGRGLFTNGS--CADIKKYEAQVYLKHLNH 467
QY 433 LTFHNG-GLPLRFSSSGNVDMEDYDLKLV---VMQGSVPLHLDYGRN-GLSKTERL----- 483
DB 468 LNTFSMGQVTFDECGDLAGNYSIIMWHLSPEDGSIV-FKEVGYYNVYAKKGERLFIN 526
QY 484 -KIRMTSDNQKPVSRCSQCOEGQVRR-VKGFHSCCYCDQCEAGSYQNPDDICTATCFC 541
DB 527 EKILNMGFRREYFSCSDICLAGTRKGIIEBPTCCFCEVCECPDCEYSDFTDASCDKIC 586
QY 542 GODEMSPERSTRCFRRRSRFLAMGEPVALLLLLSLALGLVLAALGLFVHHRDSPLVQA 601

DB 587 PDDFMSENHNTSCIAKEIFELMTBEPFGIALLLFANVLGFLAFLVGLVETKRRNPIYVA 646
QY 602 SCGPPLA---CFGLVCLGLVCL-SVLLFPQSPBARCLAOQPSHLPLTCCSLTFLQAAE 657
DB 647 TUNELSYLLLFSLIC---CFSSSLFFTEGPDDMTICRLQPAFGISFVLCISCIWTKTR 702
QY 658 IFV--ESELPLS---MADRLSGCLRGPMWLVVLAALVEVALCTWLVAFPEPVND 710
DB 703 VLVLEPAKIPTEFRHKKMGLNQ-----FLVFLCTFMQIVICATMLNTAPSSYRN 754
QY 711 WHMLPTEAL-VHCRSRWSVFGIAHATNATLAFLCLGFLVRSQPCYNNRARGLPFAML 769
DB 755 -HEDEELIFITCHEGSLMALGFLGLYTCLLAIACFFAFKSKRLPENNEAKETFSML 813
QY 770 AFTITWSFVPLLANQVYLRAVQMGALLLCLVGLIAAFHLPRCLMLRQGLNTP 827
DB 814 IFFIWIISPYPAYASTYGRFVSNAVEVIALAASFGILACIFENKVIYIILFKPSRTIE 871

RESULT 10
US-08-484-565-5
Sequence 5, Application US/08484565
Patent No. 5763569
GENERAL INFORMATION:
APPLICANT: Edward M. Brown
APPLICANT: Steven C. Hebert
APPLICANT: James E. Garrett, Jr.
TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Lyon & Lyon
STREET: First Interstate World Center
STREET: Suite 4700
STREET: 633 West Fifth Street
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90071
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: FASTSEQ
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,565
FILING DATE: 7 June, 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
PRIOR APPLICATION DATA: Including application
PRIOR APPLICATION DATA: described below: 9
APPLICATION NUMBER: 08/353,784
FILING DATE: 9 December, 1994
APPLICATION NUMBER: PCT/US/94/12117
FILING DATE: 21 October, 1994
APPLICATION NUMBER: U.S. 08/292,827
FILING DATE: 23 August, 1994
APPLICATION NUMBER: U.S. 08/141,248
FILING DATE: 22 October, 1993
APPLICATION NUMBER: U.S. 08/009,389
FILING DATE: 23 February, 1993
APPLICATION NUMBER: U.S. 08/017,127
FILING DATE: 12 February, 1993
APPLICATION NUMBER: U.S. 07/934,161
FILING DATE: 21 August, 1992
APPLICATION NUMBER: U.S. 07/834,044
FILING DATE: 11 February, 1992
APPLICATION NUMBER: U.S. 07/749,451
FILING DATE: 23 August, 1991
ATTORNEY/AGENT INFORMATION:
NAME: Heber, Sheldon O.
REGISTRATION NUMBER: 38,179

REFERENCE/DOCKET NUMBER: 213/006
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1085 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-480-565-5

Query Match 23.4%; Score 1060; DB 1; Length 1085;
 Best Local Similarity 30.0%; Pred. No. 1,2e-92;
 Matches 269; Conservative 155; Mismatches 362; Indels 112; Gaps 26;

QY 6 VLGLSLALHPGAPGLCSQQLRMKDDYVGLGLPL--GBA-EEAGLRKRTSPSSVPC 62
 DB 10 LLAFTWCTSAVGP-----DQRAQKGDIIIGLPLHFGVAVKQDPLKS--RESEVEC 61
 QY TRFSSNGLLMALMKMAVEEINNKSDLPGLRLGDEPTCSEPVVAMKPSIMFLA--KA 120
 DB 62 IRYNFRGRMIOAMFAIEEINSSPALLPNMGLTRIDTCNVSKALEATLSEVAONKI 121
 QY 121 GSRDIAVCNTYQAPRYLAVIGPSSSELAVYTKFESFPLMPOVSYGASMELLSARETF 180
 DB 122 DSLNDEFRCNCEHPISTIAYVATGSGISTAVANLGLFTIPQSVASSSSRLSNKNOF 181
 QY 181 PSFRTVSDVQVLTAAELLQERGMNVVAAAGSDDEYRGQGLSTFSLLAAARGICIAHE 240
 DB 182 KSFLETPNDHQATAMADIIIEYFRMNWVGTTAADDYGRGIEKFRREAEERDIDFS 241
 QY 241 GLVPLPRADSRIGVQDYLHVQVNSVQVLLFASVAAHAALFNYSISSRLSPVWVAS 300
 DB 242 ELI-----SQYDEEKIQVVEIQTSTAKVIVPSSGPDLEPLKEIYRNTGTIMWLAS 297
 QY 301 EAMLTSLVWGLPGMAQGVLTQGAQLHEEPQVY-KTHLALATDPAFCSAL----- 354
 DB 298 EAMASSSLIAMPDEYHVVGITIGFLKAGQIPGREFLOKVPKRSVHNGFAKEFEETFE 357
 QY 355 -----GEREGGL-----EEDVVGQRQCQCOCITON 380
 DB 358 NCHLOEAKGKPLVDFTLRGHEBEGARLSNSPTAFRPLCTGSEINSSVETPMDTHLR- 416
 QY 381 VSAGLNHHQTESVYAAYVSAQALHNTLQC-----NASGCPAQDVPKPMOLENNYN 432
 DB 417 -----ISYNYLAVYSIAHALODITYCIPGRGLFTNGS-CADIKKVEAMOVAKLHRH 467
 QY 433 LTFHYG-GLPLRFDSGGVNDMEYDILM---VMGGSVRLHDVGRFN-GSLTERL---- 483
 DB 468 LNFYSNMGEQVTFDECGGLAGNYSIINWHLSPEDGSIV-FKEVGYYNYANKGERLFTND 526
 QY 484 -KIRHNTSDNCKPVSRGROGEOGVRR-VKGFHSQCYDVCDEGSRONDDIACITFC 541
 DB 527 EKILMSGSRVPPSNCROCLAGRKGIIRGEPPCECEVCPDGEISDETASACCKC 586
 QY 542 GODEMSPERSTRCFRRRSRFLAMGEPVALLLLLSIALGVLALGIFVHHRDSPLVOA 601
 DB 587 PDDMSSNNHNSCIAKEFLFSWTEPEFGIALTFRAVLGIFLFAVLGIFIRNPIYKA 646
 QY 602 SGGPLA---CGLYVCLGLVCL-SVLLPFGQSPARCLOQPLSHLPLTGCLSTELQAAE 657
 DB 647 TNRELSTYLLTSLIC-----CFSSSLFIFGEPODWCRLRPAFGISVLCISCLIVKTR 702
 QY 658 IFV--ESLPLS-----WADRLSGCLRPMAMLVLLAMLYEVALCTVYLAAPPEVTD 710
 DB 703 VLVFEAKIPTSPHRKMWGLNQ-----FLVFLCTEQIYCAIWLMTAPSSSYRN 754
 QY 711 WHMLPTLAL-VHCRTSRVSWFGLAHATNATLALFCLEGTIFYRSOPGYNRARGLTFML 769
 DB 755 -HELEDEIIFITCHRGSLMALGLIGTYCLLAALICFPFAFNRKLPENFNRAKPTTFPSML 813

QY 770 AFTTWSEFVELLANVQVLRPAVOMGALLCVGLLAAPHLPKCYLIMRQPGINTPE 827
 DB 814 IFFIWTWSEIFAYASTGKFPVSANVEVALIALLASGLLACIFENNVYITLIFKPSNRTLE 871

RESULT 11
 US-08-480-751-5
 Sequence 5, Application US/08480751
 Patent No. 5858684
 GENERAL INFORMATION:
 APPLICANT: Edward F. Nemeth
 APPLICANT: Edward M. Brown
 APPLICANT: Steven C. Hebert
 APPLICANT: Forrest H. Fuller
 APPLICANT: James E. Garrett, Jr.
 TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 NUMBER OF INVENTION: MOLECULES
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Lyon & Lyon
 STREET: First Interstate World Center
 STREET: Suite 4700
 STREET: 633 West Fifth Street
 CITY: Los Angeles
 STATE: California
 COUNTRY: USA
 ZIP: 90071
 COMPUTER READABLE FORM:
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: FASTSEQ
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/480,751
 FILING DATE: 7 June, 1995
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 PRIOR APPLICATION DATA: including application
 PRIOR APPLICATION DATA: described below: 9
 APPLICATION NUMBER: 08/353,784
 FILING DATE: 9 December, 1994
 APPLICATION NUMBER: PCT/US/94/12117
 FILING DATE: 21 October, 1994
 APPLICATION NUMBER: U.S. 08/292,827
 FILING DATE: 23 August, 1994
 APPLICATION NUMBER: U.S. 08/141,248
 FILING DATE: 22 October, 1993
 APPLICATION NUMBER: U.S. 08/009,389
 FILING DATE: 23 February, 1993
 APPLICATION NUMBER: U.S. 08/017,127
 FILING DATE: 12 February, 1993
 APPLICATION NUMBER: U.S. 07/934,161
 FILING DATE: 21 August, 1992
 APPLICATION NUMBER: U.S. 07/834,044
 FILING DATE: 11 February, 1992
 APPLICATION NUMBER: U.S. 07/749,451
 FILING DATE: 23 August, 1991
 ATTORNEY/AGENT INFORMATION:
 NAME: Hebert, Sheldon O.
 REGISTRATION NUMBER: 38,179
 REFERENCE/DOCKET NUMBER: 213/004
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1085 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-480-751-5

Query Match 23.4%; Score 1060; DB 2; Length 1085;
 Best Local Similarity 30.0%; Pred. No. 1.2e-92;
 Matches 269; Conservative 155; Mismatches 362; Indels 112; Gaps 26;

```

QY 6 VIGLSIMALLHRTGAPLCSOOLRMKGDYVIGLPL--GEA-EBAGLRSTRPSPVC 62
DB 10 LIAFSTWCTSAVGP-----DQRAOKKGDIIIGLGFPIHFVAVVADQDIKS--RPSVVC 61
QY 63 TRFSSNGILMALAMKAVEINNKSDLLPGLRLGDLPTCSEPPVAMKPSIMFLA--KA 120
DB 62 IRYNFRGFRWLQAMIFAEIENSSPALPLNMFLGRIPTCTMTVSALATLTSFAVQNKI 121
QY 121 GSRDIAAYCNYQYQPRYLAVIGPSSSELAWTGKFFSFLLMPOYSYGSMELLSARETF 180
DB 122 DSLNDEFNCSEHIPSTIAVVGATGSGISTAVANLGLFTYPOYSYASSSLSKNOF 181
QY 181 PSFRTVPSDRVQVLTAAAEILOEFGMNVAALGSDDEXRGOLSTIFSALAAAGCICIAHE 240
DB 182 KSFLLRTIPNDEHQATAMADIIEYFRMNWGTIAADDDYGRPGIEKPREAEERDIDERS 241
QY 241 GLVPLPRADDSRLGVQDYLHOVNOSSVOYVLLFASVNAHLNFTSISRLSPKWWAS 300
DB 242 ELI-----SOYSDERKIQOYVEVIONSTAKVIVFSSGPDLEPLIKETVRNITGRIMLAS 297
QY 301 EAMLTSDIWMGLPMAOMQTVLGFLOGAQLHEFPQYV-KTHALATDPAFCSAL----- 354
DB 298 EAMASSSLIAMEYFHVVGITIGFGLKAGQIGFREFLOKVPKRSVNHGFAKEFEET 357
QY 355 -----GEREGL-----EDVVGQRCPQDCITLON 380
DB 356 NCHLOGAKGPLVDTFLNGHEGAGRLSNSPTAFRPLCTGEBNISVETPYMDYTHLR- 416
QY 381 VSAGLNHQTSEYVAAYTSVAQALNHTLOC-----NAGSCPQDPKPMOLLENMWN 432
DB 417 -----ISYVYLAIVSIALADIDITICIGRGLFTNGS-CADIKKYEAQVLAHLMH 467
QY 433 LTFHVG-GLPLRDSGNDMEYDKLW---VMQSVPLHNVGRN-GSLTERL----- 483
DB 468 LNFSTNMGEQVTFDECGDLAGNYSIINMHLSPEDSIV-FKEVGYNYVAKKGERLFIND 526
QY 484 -KIRNHTSNOKPVRCSCQCOEGOVR-VKGFHSCCTVCVCEAGSYQNPDDIACITC 541
DB 527 EKILMSGFSREVFSNCSDCIAGRKGIIEGPCCFECVBCPGEYSDETDASACDNC 586
QY 542 GODEMSPERSTRCFRRSRRLAMGPAYVLLLLLSIALGLVLAAGLVHHRDSPLVA 601
DB 587 PDEFNSNENHTCIAKEIFELTSMTEPGIALTLFAVLGIFLAFLVGIKTRNPTIYVA 646
QY 602 SGGPLA--CFGLVCLGLVCL-SVLLFPQSPARCLAQOPLSHLPLTGLSTLFLQAAE 657
DB 647 TNRRLSYLLLFSLC---CFSSSLFEGEPQDMTCRLQPAFGISFVICISCIILVKNR 702
QY 658 IIV--ESELPLS-----WADRLSGCLRGPMWLVYLAAMEVALCTWYLAAPPEVTD 710
DB 703 VLLVFEAKIPFSHRRKMWGLNQ-----FLVFLCTFMOIVICALMLNAPSSSYRN 754
QY 711 WHMLPTEAL-VHCRTRSWVSFGLAHATNATLAFELCFGLFVLRSOPGCBNRARGLEFAM 769
DB 755 -HELEDEIFITICHSGLMALGFLIGYTCLLAICFFFAFKSRKLDENFNKAFITFSML 813
QY 770 AYTITWVSFPLLANVOVLLRPVOMGALLLCVLGLAFLHPRCLYLMKROGLMTPPE 827
DB 814 IFFIWISSIPAYASTYGFVSAAVEVIALIASFGLLACIFENKYYIILFKRSRMTIE 871

```

RESULT 12
 US-08-943-986-5
 ; Sequence 5, Application US/08943986
 ; Patent No. 5962314
 ; GENERAL INFORMATION:
 ; APPLICANT: Edward M. Brown
 ; APPLICANT: Steven C. Hebert
 ; APPLICANT: James E. Garrett, Jr.

```

; TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
; TITLE OF INVENTION: MOLECULES
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: First Interstate World Center
; STREET: Suite 4700
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: USA
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: FASTSEQ
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/943,986
; FILING DATE: 03-OCT-1997
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/484,565
; FILING DATE: 7-June-1995
; APPLICATION NUMBER: 08/353,784
; FILING DATE: 9 December, 1994
; APPLICATION NUMBER: PCT/US/94/12117
; FILING DATE: 21 October, 1994
; APPLICATION NUMBER: U.S. 08/292,827
; FILING DATE: 23 August, 1994
; APPLICATION NUMBER: U.S. 08/141,248
; FILING DATE: 22 October, 1993
; APPLICATION NUMBER: U.S. 08/009,389
; FILING DATE: 23 February, 1993
; APPLICATION NUMBER: U.S. 08/017,127
; FILING DATE: 12 February, 1993
; APPLICATION NUMBER: U.S. 07/934,161
; FILING DATE: 21 August, 1992
; APPLICATION NUMBER: U.S. 07/834,044
; FILING DATE: 11 February, 1992
; APPLICATION NUMBER: U.S. 07/749,451
; FILING DATE: 23 August, 1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Heber, Sheldon O.
; REGISTRATION NUMBER: 38,179
; REFERENCE/DOCKET NUMBER: 213/006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELE: 67-3510
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1085 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-943-986-5

```

Query Match 23.4%; Score 1060; DB 2; Length 1085;
 Best Local Similarity 30.0%; Pred. No. 1.2e-92;
 Matches 269; Conservative 155; Mismatches 362; Indels 112; Gaps 26;

```

QY 6 VIGLSIMALLHRTGAPLCSOOLRMKGDYVIGLPL--GEA-EBAGLRSTRPSPVC 62
DB 10 LIAFSTWCTSAVGP-----DQRAOKKGDIIIGLGFPIHFVAVVADQDIKS--RPSVVC 61
QY 63 TRFSSNGILMALAMKAVEINNKSDLLPGLRLGDLPTCSEPPVAMKPSIMFLA--KA 120
DB 62 IRYNFRGFRWLQAMIFAEIENSSPALPLNMFLGRIPTCTMTVSALATLTSFAVQNKI 121
QY 121 GSRDIAAYCNYQYQPRYLAVIGPSSSELAWTGKFFSFLLMPOYSYGSMELLSARETF 180
DB 122 DSLNDEFNCSEHIPSTIAVVGATGSGISTAVANLGLFTYPOYSYASSSLSKNOF 181

```


Db 358 NCHLOEAGAKGPLVDITLHGHEGAGRLSNSPTAFRLPCTGEBNISVETPYMDYTHLR- 416
 QY 381 VSAGLNHQTSTVAAYVSAQALHNTLOC-----NASGCPADDPVKPQOLLENNYN 432
 Db 417 -----ISYNYLVAYVSAIALADITYCTIGRGLFTNGS-CADIKKVEAMQVYKHLRH 467
 QY 433 LTFHVG-GLPLRFDSGGVNDMEYDKLW---VMQGSVPRLDVGRFN-GSLRTERL---- 483
 Db 468 LNTSNMGEQVTFDECGDLAGNYSTIINMHLSPEDGSIV-FKEVGYNYVAKKGERLEIND 526
 QY 484 -KIRNHTSDNOKPVSRCSQCOEGOVRR-VKGFHSCDYCDVCEAGSYRONDDIACFTG 541
 Db 527 EKILMSGFREVFPNSCRDCLAGTRKGIIEBEPCCFCEVCPCPGEXSDETDASACDCK 586
 QY 542 CODESPSPSRTRCFRRRRTFLAMGEPVALLLLSLALGLVLAALGLFVHHRSDPIVQA 601
 Db 567 PDDEFSNENHTSCIAKEIEFESWEPFGIALTFEVLGIFLTAFLVGFIFERNPIYKA 646
 QY 602 SGGPLA--CFGLVCLGLVCL-SVLLFPQSPPARCLAQDPLSHLPLTGLSTLEFLQAAE 657
 Db 647 TNRELSTYLLFLSLC-----CFSSSLFFIGEPQDWTCLRQPAFGISFVLCISCLIVKRN 702
 QY 658 IFV--ESLEPLS-----MADRISGCLRGPMAMLVLLMALVEVALCTWYLVAPPEVTD 710
 Db 703 VLLVEAKIPTSFHKKWGLNIQ-----FLVFLCTFMQIVYCAIMLTAPSSSYRN 754
 QY 711 WHMLPTEAL-VHCRTSRVSVFGLAHATNATLAFCELTFLVRSQPGVYNARGLTFEML 769
 Db 755 -HELFDELITFTCHGSLMALGFLGYCTCLAAICFFFAFRSKLPENENKKTFFESML 813
 QY 770 AEFITVSEVPLLAQVVLPRVOMGALLLVGLILAFAHLPRCYLLMRQGLTPE 827
 Db 814 IFFIWIWFIPAYASTGYKFSVAVEVIALIASFGLACIFENKYYIILFKRSNTIE 871

RESULT 14
 US-08-484-719B-5
 : Sequence 5, Application US/08484719B
 : Patent No. 6031003
 : GENERAL INFORMATION:
 : APPLICANT: Edward F. Nemeth, Edward M.
 : APPLICANT: Brown, Steven C. Hebert,
 : APPLICANT: Bradford C. Van Wagonen,
 : APPLICANT: Manuel F. Balandrin,
 : APPLICANT: Forrest H. Fuller, Eric G.
 : APPLICANT: Delmar, Scott T. Moe
 : TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE
 : NUMBER OF SEQUENCES: 20
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Lyon & Lyon
 : STREET: First Interstate World Center
 : STREET: Suite 4700
 : STREET: 633 West Fifth Street
 : CITY: Los Angeles
 : STATE: California
 : COUNTRY: USA
 : ZIP: 90071
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
 : OPERATING SYSTEM: MS Word
 : SOFTWARE: FASTSEQ for Windows Version 3.0
 : CURRENT APPLICATION DATA:
 : APPLICATION NUMBER: US/08/484,719B
 : FILING DATE: 7 June, 1995
 : CLASSIFICATION: 514
 : PRIOR APPLICATION DATA:
 : APPLICATION NUMBER: 08/353,784
 : FILING DATE: 9 December, 1994
 : APPLICATION NUMBER: PCT/US/94/12117
 : FILING DATE: 21 October, 1994
 : APPLICATION NUMBER: U.S. 08/292,827

: FILING DATE: 23 August, 1994
 : APPLICATION NUMBER: U.S. 08/141,248
 : FILING DATE: 22 October, 1993
 : APPLICATION NUMBER: U.S. 08/009,389
 : FILING DATE: 23 February, 1993
 : APPLICATION NUMBER: U.S. 08/017,127
 : FILING DATE: 12 February, 1993
 : APPLICATION NUMBER: U.S. 07/934,161
 : FILING DATE: 21 August, 1992
 : APPLICATION NUMBER: U.S. 07/834,044
 : FILING DATE: 11 February, 1992
 : APPLICATION NUMBER: U.S. 07/749,451
 : FILING DATE: 23 August, 1991
 : ATTORNEY/AGENT INFORMATION:
 : NAME: Douglas C. Murdock
 : REGISTRATION NUMBER: 37,549
 : REFERENCE/DOCKET NUMBER: 213/007
 : TELECOMMUNICATION INFORMATION:
 : TELEPHONE: (213) 489-1600
 : TELEFAX: (213) 955-0440
 : Telex: 67-3510
 : INFORMATION FOR SEQ ID NO: 5:
 : SEQUENCE CHARACTERISTICS:
 : LENGTH: 1085 amino acids
 : TYPE: amino acid
 : STRANDEDNESS: single
 : TOPOLOGY: linear
 : MOLECULE TYPE: protein
 : US-08-484-719B-5

Query Match 23.4%; Score 1060; DB 3; Length 1085;
 Best Local Similarity 30.0%; Pred. No. 1.2e-92;
 Matches 269; Conservative 155; Mismatches 362; Indels 112; Gaps 26;

QY 6 VLGLSTALHLHPGTCAPLCLSQQLRMKGDYVIGLFPPL--GEA-EEAGLRSTRSPSYVC 62
 Db 10 LLAFTWCTSAVGP-----DQRAOKGDIILGFLPFIHGVAVXDOLKS--RESEVCC 61
 QY 63 TRFSSNGLLMAMKMAVEEINNKSDLLPGLRLGYDLDTGCEPVAVAKPSLMFLA--KA 120
 Db 62 IRYNRGRFWQAMLFALIEINSSPALLPNNLTGRIPTDCTVSKALAEATLSEVAKNT 121
 QY 121 GSRDIAAYCNTQYQPRVLAIVIGPSSSELAWTGKFFSFLMPQVSYGASMELSARETF 180
 Db 122 DSNLDFCNCSEHPISTIAVVGATGSGISTAVANMLGLFYIPQVSYASSSLLENKNGF 181
 QY 181 PSFFRTVSDRYQLTAABELLOEFGMNVAALGSDDEYRGGLSTFSALAAAGCICIAHE 240
 Db 182 KSFLRTIENDEHOATAMADIIIEYFWMNVGRTIADDDYGRGPIEKFRERAEERDIDIFS 241
 QY 241 GLVPLPRADDSRLGKRVODVLRHVNQSVQVLLFASVNAHALFNYSISSRLSPRYVWAS 300
 Db 242 ELI-----SOYSPDEKIQOVVEYIONSIAKVIYVFSSGPLEPLIEIYARNITGRILNAS 297
 QY 301 EAMLTSDLVMLGPGMAOMGTVGLQRGALHEFPQVY-KTHLALATPAFCSAL----- 354
 Db 298 EAMASSLIAMPEYFHVVGRTIGFGIKAGQIGRFRELOKVPKRSVHNGFAKEWEETE 357
 QY 355 -----GEREGGL-----BEDVVGQRCPQDCDITLON 380
 Db 358 NCHLOEAGAKGPLVDITLHGHEGAGRLSNSPTAFRLPCTGEBNISVETPYMDYTHLR- 416
 QY 381 VSAGLNHQTSTVAAYVSAQALHNTLOC-----NASGCPADDPVKPQOLLENNYN 432
 Db 417 -----ISYNYLVAYVSAIALADITYCTIGRGLFTNGS-CADIKKVEAMQVYKHLRH 467
 QY 433 LTFHVG-GLPLRFDSGGVNDMEYDKLW---VMQGSVPRLDVGRFN-GSLRTERL---- 483
 Db 468 LNTSNMGEQVTFDECGDLAGNYSTIINMHLSPEDGSIV-FKEVGYNYVAKKGERLEIND 526
 QY 484 -KIRNHTSDNOKPVSRCSQCOEGOVRR-VKGFHSCDYCDVCEAGSYRONDDIACFTG 541
 Db 527 EKILMSGFREVFPNSCRDCLAGTRKGIIEBEPCCFCEVCPCPGEXSDETDASACDCK 586

QY 542 GODEMSPERSTRERRSRFLANGEPVALLLLLSIALGLVLAALGLFVHHDRSPVQA 601
 Db 587 PDDFMSENHNSCIAKEIEFLSWTEPEGIALTEFAVIGIFLFAVIGIFENRNTPIKA 646
 QY 602 SGGPLA---CGVLGLVCL-SVLEFPGQSPARCQAQPLSHLPGLTGLSTFLQAAE 657
 Db 647 TNRBLSTLLFLSLC---CFSSSLFFGEPQDWTCLRPAGISFVLCISCLIVKTR 702
 QY 658 IFV--ESELPLS-----WADRLSGCLRGPMAMVLLAMVEVALCWTYLAAPPEVTD 710
 Db 703 VLVLEFAKIPTRSHRKMGLQ-----FLVFLCTFMQIVICAILMTATPSSYRN 754
 QY 711 WHMLPTEAL-VHCRTRSWVSGLAHAATNATLAEFLCTGLTVRSOPGCYNRARGITFAML 769
 Db 755 -HELEDELIIFTCHEGSLMALGFLIGYTCILAAICFFPAFRSKRLPENFNKAFITPSML 813
 QY 770 AYTFTWSPVPLANNVOVLRPAVOMGALLCVGILAAFLPRLPCYILMRPGLINTPE 827
 Db 814 IFTVWISFLPAYASTYGRKFSVAVEVIALAASFLACIPFNKYIILFRPSRNTIE 871

RESULT 15

US-08-484-159-5

Sequence 5, Application US/08484159

Patent No. 6313146

GENERAL INFORMATION:

APPLICANT: Bradford C. Van Wagenen

APPLICANT: Manuel F. Balandrin

APPLICANT: Eric G. Del Mar

TITLE OF INVENTION: CALCIUM RECEPTOR-ACTIVE

TITLE OF INVENTION: MOLECULES

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Lyon & Lyon

STREET: First Interstate World Center

STREET: Suite 4700

STREET: 633 West Fifth Street

CITY: Los Angeles

STATE: California

COUNTRY: USA

ZIP: 90071

COMPUTER READABLE FORM:

MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: FASTSEQ

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/484,159

FILING DATE: 7 June, 1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA: including application

Prior Application Data: described below: 9

APPLICATION NUMBER: 08/353,784

FILING DATE: 9 December, 1994

APPLICATION NUMBER: PCT/US/94/12117

FILING DATE: 21 October, 1994

APPLICATION NUMBER: U.S. 08/292,827

FILING DATE: 23 August, 1994

APPLICATION NUMBER: U.S. 08/141,248

FILING DATE: 22 October, 1993

APPLICATION NUMBER: U.S. 08/009,389

FILING DATE: 23 February, 1993

APPLICATION NUMBER: U.S. 08/017,127

FILING DATE: 12 February, 1993

APPLICATION NUMBER: U.S. 07/934,161

FILING DATE: 21 August, 1992

APPLICATION NUMBER: U.S. 07/834,044

FILING DATE: 11 February, 1992

APPLICATION NUMBER: U.S. 07/749,451

FILING DATE: 23 August, 1991

ATTORNEY/AGENT INFORMATION:
 NAME: Heber, Sheldon O.
 REGISTRATION NUMBER: 38,179
 REFERENCE/DOCKET NUMBER: 214/101
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (213) 489-1600
 TELEFAX: (213) 955-0440
 TELEX: 67-3510
 INFORMATION FOR SEQ ID NO: 5:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1085 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-484-159-5

Query Match 23.4%; Score 1060; DB 4; Length 1085;

Best Local Similarly 30.0%; Pred. No. 1,2e-92;

Matches 269; Conservative 155; Mismatches 362; Indels 112; Gaps 26;

QY 6 VLGSLMALHPGTGAPLCLSQOLRMKGVDVYLGGLFPL--GEA-EEAGLSRTRSPSPVC 62
 Db 10 ILASTWCTSAKGP-----DQRAQKGDITLGLFPIHFGVAVKDDOLAS--RPESVEC 61
 QY 63 TRFSSNGILMALAMKMAVEEINNKSDLLPGLRLGYDLEDFDCSEFVPMKPSIMFLA--KA 120
 Db 62 IRYFRGFRMIOAMIFAEIEINSSPALLPNTLGRIFPDICNTYSKALENTLSFAQNKI 121
 QY 121 GSRDIAVCNTQYQPRVLAIVGPHSELAVNTGKFSFLMPQVSYGASHLLSARETF 180
 Db 122 DSLMIDFCNCSEHIPSTIVAVGATGSGISTAVANLGLFVYIPQVSYASSRLLSNKNOF 181
 QY 181 PSFRTVPSDRVOLTAAEILQEFGMNVAALGSDDEYGRGGLSIFALAAARGICIAHE 240
 Db 182 KSLRTIPNDENQATAADIIETFRMNVGTIADDDYGRGICFKFEKPEBERDIDCDFS 241
 QY 241 GLVPLPRADSRIGKVODVLHQVNOSSVOYVLLFASVHAALFNYSISRLSPKVVAS 300
 Db 242 ELI-----SQXSDEERIQOQVEVIONSTAKYIVVSSGPDLEPLKEIVRRNIRIWLAS 297
 QY 301 EAMLTSLVNGILPGMAOMGVTLQGAQLHPRQV--KTHLALADPACSL----- 354
 Db 298 EAMASSSLIAMPYEFHVAGGTIGGLAKAGQIPGRFELQVHPKSVHNGFAEWEETP 357
 QY 355 -----GEREGGL-----EEDVYQRCPOCCITLON 380
 Db 358 NCHLOEAKGPLPDFTFLRHBEGGARLSPTAFRLCTGEEINISVEPMDYTLR- 416
 QY 381 VSAGLNHQTFFSYAAYVAQALHNTLOC-----NASGCPADDPKRPWOLLENMYN 432
 Db 417 -----ISYNYLAVYSIAHLADITYCTIPGRGLFTNGS-CADIKKVAEOMVLKHLRH 467
 QY 433 LTFHVG-GLPLRPDSSGNVMEYDLKLM---VMQGSVRLHDHGRFN-GSLRTRRL----- 483
 Db 468 LNTFSMGEQVYTDDECDLGNSTIIMWHLSPEDGSTIV-FKEGYIVNYAKKGERLEIND 526
 QY 484 -KIRWHTSDNQKPYSRCSROCOEQVYR--VKGFHSCCYDCEVDEAGSYRONPPDIACTFC 541
 Db 527 EKILMSGFSREVEFPNSRCLAGTRKGLIEBETCFCEVECPDGETYSETAASACDKC 586
 QY 542 GODEMSPERSTRERRSRFLANGEPVALLLLLSIALGLVLAALGLFVHHDRSPVQA 601
 Db 587 PDDFMSENHNSCIAKEIEFLSWTEPEGIALTEFAVIGIFLFAVIGIFENRNTPIKA 646
 QY 602 SGGPLA---CGVLGLVCL-SVLEFPGQSPARCQAQPLSHLPGLTGLSTFLQAAE 657
 Db 647 TNRBLSTLLFLSLC---CFSSSLFFGEPQDWTCLRPAGISFVLCISCLIVKTR 702
 QY 658 IFV--ESELPLS-----WADRLSGCLRGPMAMVLLAMVEVALCWTYLAAPPEVTD 710
 Db 703 VLVLEFAKIPTRSHRKMGLQ-----FLVFLCTFMQIVICAILMTATPSSYRN 754
 QY 711 WHMLPTEAL-VHCRTRSWVSGLAHAATNATLAEFLCTGLTVRSOPGCYNRARGITFAML 769

Db 755 -HELEDEIIFITCHEGSLMALGFLIGYTCLLAACEFFAFKSRKLPENFNEAKFITFSML 813
QY 770 AYEITWSEFVPLANQVVLAPAVOMGAILLCVLGIIAFLHLPKCYILMRQPLNTPPE 827
Db 814 IFFIWMISFIPAYASTYGFVSAVEYAIALLASFGILACIFENKVIYIILFKPSRNTTE 871

Search completed: May 19, 2003, 09:53:24
Job time : 27.6201 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.4.p5.4578
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 19, 2003, 09:51:25 ; Search time 39.3231 seconds
(without alignments)
2089.767 Million cell updates/sec

Title: US-09-927-315-15

Perfect score: 4524
Sequence: 1 MGPAVLGLSIMALHPGTG.....GPGDAQQNDGNTGNCKRHE 852

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 362588 seqs, 96450795 residues

Total number of hits satisfying chosen parameters: 362588

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published_Applications_AA:*

1: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB pep:*\n2: /cgn2_6/ptodata/2/pubpaa/PCR_NEW_PUB pep:*\n3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB pep:*\n4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB pep:*\n5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB pep:*\n6: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB pep:*\n7: /cgn2_6/ptodata/2/pubpaa/PCRUS_PUBCOMB pep:*\n8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB pep:*\n9: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB pep:*\n10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB pep:*\n11: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB pep:*\n12: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB pep:*\n13: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB pep:*\n14: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	4524	100.0	852	9	US-09-927-427A-6
2	4524	100.0	852	9	US-09-927-315-15
3	4524	100.0	852	9	US-10-035-045-4
4	4512	99.7	852	9	US-10-261-482-2
5	4512	99.7	852	9	US-10-287-837-14
6	4512	99.7	852	10	US-09-796-338A-14
7	4500	99.5	850	9	US-09-799-629-4
8	3255	71.9	858	9	US-09-799-629-14
9	3255	71.9	858	9	US-09-927-315-25
10	3255	71.9	858	9	US-10-035-045-14
11	3229	71.4	858	9	US-09-927-315-23
12	3227.5	71.3	858	9	US-09-927-315-18
13	3221.5	71.2	858	9	US-09-927-315-20
14	1769	39.1	1138	9	US-10-261-482-4
15	1208.5	26.7	840	9	US-09-361-652-1
16	1208.5	26.7	840	9	US-09-927-315-1
17	1188.5	26.3	842	9	US-09-361-652-2
18	1188.5	26.3	842	9	US-09-927-315-2
19	1153	25.3	841	9	US-09-897-427A-2

20	1153	25.5	841	9	US-09-799-629-17	Sequence 17, Appl
21	1153	25.5	841	9	US-10-035-045-17	Sequence 17, Appl
22	1152	25.5	841	10	US-09-819-946-2	Sequence 2, Appl
23	1082.5	23.9	763	10	US-09-819-946-4	Sequence 4, Appl
24	1068.5	23.6	1078	10	US-09-727-205-2	Sequence 2, Appl
25	1068.5	23.6	1078	12	US-10-002-854-2	Sequence 2, Appl
26	1067.5	23.6	1078	9	US-10-125-792-28	Sequence 28, Appl
27	1067.5	23.6	1078	9	US-10-125-778-28	Sequence 28, Appl
28	1065	23.5	941	9	US-10-125-792-8	Sequence 8, Appl
29	1065	23.5	941	9	US-10-125-778-8	Sequence 8, Appl
30	1064	23.5	941	9	US-10-125-792-10	Sequence 10, Appl
31	1064	23.5	941	9	US-10-125-778-10	Sequence 10, Appl
32	1056	23.3	1027	9	US-10-125-792-2	Sequence 2, Appl
33	1056	23.3	1027	9	US-10-125-778-2	Sequence 2, Appl
34	1053	23.3	777	9	US-09-361-652-3	Sequence 3, Appl
35	1053	23.3	777	9	US-09-927-315-3	Sequence 3, Appl
36	1048.5	23.2	850	9	US-10-125-792-12	Sequence 12, Appl
37	1048.5	23.2	850	9	US-10-125-778-12	Sequence 12, Appl
38	1027.5	22.7	926	10	US-09-816-685-2	Sequence 4, Appl
39	1014.5	22.4	839	9	US-09-897-427A-4	Sequence 2, Appl
40	1014.5	22.4	839	9	US-10-035-045-21	Sequence 21, Appl
41	1011	22.3	838	9	US-09-927-315-9	Sequence 9, Appl
42	1008.5	22.3	828	10	US-09-816-685-4	Sequence 4, Appl
43	1002.5	22.2	843	9	US-10-096-144-1	Sequence 1, Appl
44	1002.5	22.2	843	9	US-09-927-315-7	Sequence 7, Appl
45	1002.5	22.2	843	12	US-10-124-598-1	Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-09-897-427A-6
; Sequence 6, Application US/09897427A
; Patent No. US20020160424A1
; GENERAL INFORMATION:
; APPLICANT: ADLER, JON ELLIOT
; APPLICANT: LI, XIADONG
; APPLICANT: STALEWSKI, LENA
; APPLICANT: XU HONG
; APPLICANT: EHEVERRI, FERNANDO
; TITLE OF INVENTION: T1R HETERO-OLIGOMERIC TASTE RECEPTORS
; FILE REFERENCE: 078003-0282558
; CURRENT APPLICATION NUMBER: US/09/897, 427A
; CURRENT FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 852
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-897-427A-6

Query Match	Best Local Similarity	Score	DB 9;	Length	852;
Matches 852; Conservative	100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0;	Gaps 0;
1	MGPAVLGLSIMALHPGTGAPLCISQOLRMKGDVVLGGLPFGRAEAGLRSPSP	60			
1	MGPAVLGLSIMALHPGTGAPLCISQOLRMKGDVVLGGLPFGRAEAGLRSPSP	60			
61	VCTRSSNGLMALMKMAVEEINKKSDLLPLRLGYDLFDTCSPVYAMKPSLFLAKA	120			
61	VCTRSSNGLMALMKMAVEEINKKSDLLPLRLGYDLFDTCSPVYAMKPSLFLAKA	120			
121	GSRDIAVCNTYQYPRVAVIGPSSSEIAYWGKFFSFFLMPQVSYGASMEIARSRETF	180			
121	GSRDIAVCNTYQYPRVAVIGPSSSEIAYWGKFFSFFLMPQVSYGASMEIARSRETF	180			
161	PSFFFTVSDRQVLAELAELQEFGRNMVYALGSDDEYGRGSLTFSLAARGICIAHE	240			
161	PSFFFTVSDRQVLAELAELQEFGRNMVYALGSDDEYGRGSLTFSLAARGICIAHE	240			

```

QY 241 GLVPLPRADSRRLGKVDVYLHQVNOSSVQVYLLFASVHAHALNYSISSLSRPRVWAS 300
DB 241 GLVPLPRADSRRLGKVDVYLHQVNOSSVQVYLLFASVHAHALNYSISSLSRPRVWAS 300
QY 301 EAMLTSDIYMLGPGMAOMGTVLGFLQKGAOLHEFPQYKTHLATADPAFCALSAGEROG 360
DB 301 EAMLTSDIYMLGPGMAOMGTVLGFLQKGAOLHEFPQYKTHLATADPAFCALSAGEROG 360
QY 361 LEEDVVGRCPOCCITLQNVASAGLNHHQTFSVYAAYVSAVALHNTLQCNASGCPADP 420
DB 361 LEEDVVGRCPOCCITLQNVASAGLNHHQTFSVYAAYVSAVALHNTLQCNASGCPADP 420
QY 421 VKPQOLLENNMTLTFHHVGGFLPLRFDSGNDMEYDLKLMWQGSVPRLDHVGRENGSLRT 480
DB 421 VKPQOLLENNMTLTFHHVGGFLPLRFDSGNDMEYDLKLMWQGSVPRLDHVGRENGSLRT 480
QY 481 ERLKIRMTSDNOKPVSRCSQCEGOVRRYKGFHSCCYDCVDEAGSYRONPDIACTF 540
DB 481 ERLKIRMTSDNOKPVSRCSQCEGOVRRYKGFHSCCYDCVDEAGSYRONPDIACTF 540
QY 541 CGODEWSPERSTRCFRRRSRFLANGEPAYLLLLLSLATGLVLAALGLFVHHRDSPLVQ 600
DB 541 CGODEWSPERSTRCFRRRSRFLANGEPAYLLLLLSLATGLVLAALGLFVHHRDSPLVQ 600
QY 601 ASGGPLACFGIYVCLGIVCLSVLTFPGQSPARCIAQOPLSHLPTGCLSTFLQAAETV 660
DB 601 ASGGPLACFGIYVCLGIVCLSVLTFPGQSPARCIAQOPLSHLPTGCLSTFLQAAETV 660
QY 661 ESELPLSWADRLSGCLNGPMAMLVLLAMLEVALCTWYLAFFPEVYTDHMLPTETLV 720
DB 661 ESELPLSWADRLSGCLNGPMAMLVLLAMLEVALCTWYLAFFPEVYTDHMLPTETLV 720
QY 721 HCRTRSMVSEGLAHATNATLAFICFLGTFLVRSQPGCYNRARGLTFAMLAFTWSEVP 780
DB 721 HCRTRSMVSEGLAHATNATLAFICFLGTFLVRSQPGCYNRARGLTFAMLAFTWSEVP 780
QY 781 LLANVOVYLRAVOMGALLLCVIGTILAAFHLPKCYLLMRQGLNTPPEFLGGGPDAGQ 840
DB 781 LLANVOVYLRAVOMGALLLCVIGTILAAFHLPKCYLLMRQGLNTPPEFLGGGPDAGQ 840
QY 841 NDGNTGNQKHE 852
DB 841 NDGNTGNQKHE 852

```

RESULT 2

```

; Sequence 15, Application US/09927315
; Publication No. US20030040045A1
; GENERAL INFORMATION:
; APPLICANT: Zuker, Charles S.
; APPLICANT: Ryba, Nicholas J.P.
; APPLICANT: Nelson, Greg
; APPLICANT: Hoon, Mark A.
; APPLICANT: Chandrasekar, Jayaram
; APPLICANT: Zhang, Yifeng
; APPLICANT: The Regents of the University of California
; APPLICANT: The Government of the United States of America
; APPLICANT: as represented by the Secretary of the
; TITLE OF INVENTION: Mammalian Sweet Taste Receptors
; FILE REFERENCE: 02307E-12011005
; CURRENT APPLICATION NUMBER: US/09/927,315
; PRIOR APPLICATION NUMBER: US 60/302,898
; PRIOR FILING DATE: 2001-07-03
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 15
; LENGTH: 852
; TYPE: PRM
; ORGANISM: Homo sapiens
; FEATURE:

```

```

; OTHER INFORMATION: human T1R3 sweet taste receptor
US-09-927-315-15

```

```

Query Match 100.0%; Score 4524; DB 9; Length 852;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 852; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MLGPAVLGLSLMALHGTGAPLCLSOOLRMKGDYVGLGLEPLGEAEAGLRSTRSPSP 60
DB 1 MLGPAVLGLSLMALHGTGAPLCLSOOLRMKGDYVGLGLEPLGEAEAGLRSTRSPSP 60
QY 61 VCTRFSSNGLLMALAMKMAVEINNKSDLLPLRLGIDLEDTCEBPVYAMKPSIMFLAKA 120
DB 61 VCTRFSSNGLLMALAMKMAVEINNKSDLLPLRLGIDLEDTCEBPVYAMKPSIMFLAKA 120
QY 121 GSRDIAYCANTQOYOPRLAYIGPHSSSLAMVTKFFSFFLMPQVSGASKELLISARETF 180
DB 121 GSRDIAYCANTQOYOPRLAYIGPHSSSLAMVTKFFSFFLMPQVSGASKELLISARETF 180
QY 181 PSFRTVPDSRVOQLTAAAEILLQEEFGMMVVAALGSDDEYGRQGLSIFSALAAAGICIAHE 240
DB 181 PSFRTVPDSRVOQLTAAAEILLQEEFGMMVVAALGSDDEYGRQGLSIFSALAAAGICIAHE 240
QY 241 GLVPLPRADSRRLGKVDVYLHQVNOSSVQVYLLFASVHAHALNYSISSLSRPRVWAS 300
DB 241 GLVPLPRADSRRLGKVDVYLHQVNOSSVQVYLLFASVHAHALNYSISSLSRPRVWAS 300
QY 301 EAMLTSDIYMLGPGMAOMGTVLGFLQKGAOLHEFPQYKTHLATADPAFCALSAGEROG 360
DB 301 EAMLTSDIYMLGPGMAOMGTVLGFLQKGAOLHEFPQYKTHLATADPAFCALSAGEROG 360
QY 361 LEEDVVGRCPOCCITLQNVASAGLNHHQTFSVYAAYVSAVALHNTLQCNASGCPADP 420
DB 361 LEEDVVGRCPOCCITLQNVASAGLNHHQTFSVYAAYVSAVALHNTLQCNASGCPADP 420
QY 421 VKPQOLLENNMTLTFHHVGGFLPLRFDSGNDMEYDLKLMWQGSVPRLDHVGRENGSLRT 480
DB 421 VKPQOLLENNMTLTFHHVGGFLPLRFDSGNDMEYDLKLMWQGSVPRLDHVGRENGSLRT 480
QY 481 ERLKIRMTSDNOKPVSRCSQCEGOVRRYKGFHSCCYDCVDEAGSYRONPDIACTF 540
DB 481 ERLKIRMTSDNOKPVSRCSQCEGOVRRYKGFHSCCYDCVDEAGSYRONPDIACTF 540
QY 541 CGODEWSPERSTRCFRRRSRFLANGEPAYLLLLLSLATGLVLAALGLFVHHRDSPLVQ 600
DB 541 CGODEWSPERSTRCFRRRSRFLANGEPAYLLLLLSLATGLVLAALGLFVHHRDSPLVQ 600
QY 601 ASGGPLACFGIYVCLGIVCLSVLTFPGQSPARCIAQOPLSHLPTGCLSTFLQAAETV 660
DB 601 ASGGPLACFGIYVCLGIVCLSVLTFPGQSPARCIAQOPLSHLPTGCLSTFLQAAETV 660
QY 661 ESELPLSWADRLSGCLNGPMAMLVLLAMLEVALCTWYLAFFPEVYTDHMLPTETLV 720
DB 661 ESELPLSWADRLSGCLNGPMAMLVLLAMLEVALCTWYLAFFPEVYTDHMLPTETLV 720
QY 721 HCRTRSMVSEGLAHATNATLAFICFLGTFLVRSQPGCYNRARGLTFAMLAFTWSEVP 780
DB 721 HCRTRSMVSEGLAHATNATLAFICFLGTFLVRSQPGCYNRARGLTFAMLAFTWSEVP 780
QY 781 LLANVOVYLRAVOMGALLLCVIGTILAAFHLPKCYLLMRQGLNTPPEFLGGGPDAGQ 840
DB 781 LLANVOVYLRAVOMGALLLCVIGTILAAFHLPKCYLLMRQGLNTPPEFLGGGPDAGQ 840
QY 841 NDGNTGNQKHE 852
DB 841 NDGNTGNQKHE 852

```

RESULT 3

```

; Sequence 4, Application US/10035045
; Publication No. US20030054448A1
; GENERAL INFORMATION:

```

```

; APPLICANT: ADLER, JON ELLIOT
; APPLICANT: LI, XIAODONG
; APPLICANT: STASZEWSKI, LENA
; APPLICANT: O'CONNELL, SHAWN
; APPLICANT: ZOZOLYA, SERGEY
; TITLE OF INVENTION: T1R TASTE RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: 078003-0280681
; CURRENT APPLICATION NUMBER: US/10/035,045
; CURRENT FILING DATE: 2002-01-03
; PRIOR APPLICATION NUMBER: 60/259,227
; PRIOR FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: 60/284,547
; PRIOR FILING DATE: 2001-04-19
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 852
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-035-045-4

```

```

Query Match      100.0%; Score 4524; DB 9; Length 852;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 852; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MCGPAVLGSLTMAALLHPGTGAPCLISQQLRMKGDYVIGLFLPGAEEAGLSRTRPSSP 60
DB 1 MCGPAVLGSLTMAALLHPGTGAPCLISQQLRMKGDYVIGLFLPGAEEAGLSRTRPSSP 60
QY 61 VCTRESSNGLTMAALMKAAVEEINNKSDLLPGIRLGIDYDFDCSEPVVAMKPSLMFLAKA 120
DB 61 VCTRESSNGLTMAALMKAAVEEINNKSDLLPGIRLGIDYDFDCSEPVVAMKPSLMFLAKA 120
QY 121 GSRDIAAYCNYOYOPRYAVIGPHSSELAAMYTGKFFSEFLMPQVSYGSMELLSARETF 180
DB 121 GSRDIAAYCNYOYOPRYAVIGPHSSELAAMYTGKFFSEFLMPQVSYGSMELLSARETF 180
QY 181 PSFFRTVPDRVQVTAALAEELQEFQWNVVAAALGSDDEYGRGISTIFSALAAAGCICIAHE 240
DB 181 PSFFRTVPDRVQVTAALAEELQEFQWNVVAAALGSDDEYGRGISTIFSALAAAGCICIAHE 240
QY 241 GLVPLPRADDSRLGKYQVDLHOVNOSSVOYVLLFASVHAHAHLFNTSISRSPKXWVAS 300
DB 241 GLVPLPRADDSRLGKYQVDLHOVNOSSVOYVLLFASVHAHAHLFNTSISRSPKXWVAS 300
QY 301 EAMLTSDLVMGLPMAOMGTVLGFLORGAQLHEFPQVYVTHLALATDPAFCALGEREG 360
DB 301 EAMLTSDLVMGLPMAOMGTVLGFLORGAQLHEFPQVYVTHLALATDPAFCALGEREG 360
QY 361 LEEDVVGQRCPQDCITLTONVSAGLNHQTFSVYAAVYSVAQALNHTLQCNASGCPADP 420
DB 361 LEEDVVGQRCPQDCITLTONVSAGLNHQTFSVYAAVYSVAQALNHTLQCNASGCPADP 420
QY 421 VAPWOLLEMMYLTFRVHGGLPLRFDSGVDMEYDKLVMWOGSVPRLDHVRFNGLRT 480
DB 421 VAPWOLLEMMYLTFRVHGGLPLRFDSGVDMEYDKLVMWOGSVPRLDHVRFNGLRT 480
QY 481 EELKTRMHNSDNOKPVSRCROCOBQVRRVKGFSHCCVDCVCEGSTRONPDIDACTF 540
DB 481 EELKTRMHNSDNOKPVSRCROCOBQVRRVKGFSHCCVDCVCEGSTRONPDIDACTF 540
QY 541 CGODEMSPSRSTRCFRRRSRFLAMGEPAYLILLLILSLALGIVLALGILFVHHRSPLVQ 600
DB 541 CGODEMSPSRSTRCFRRRSRFLAMGEPAYLILLLILSLALGIVLALGILFVHHRSPLVQ 600
QY 601 ASGGPLACGLVCLGIVCLSVLLFPGQSPARCLAQOPLSHLPGLGCLSTLELQAAEIV 660
DB 601 ASGGPLACGLVCLGIVCLSVLLFPGQSPARCLAQOPLSHLPGLGCLSTLELQAAEIV 660
QY 661 ESELPLSNADRLSGCLRGMAWLVYLLAEVALCTWTVAFPPEVYVDWMLPREALY 720
DB 661 ESELPLSNADRLSGCLRGMAWLVYLLAEVALCTWTVAFPPEVYVDWMLPREALY 720

```

```

QY 721 HCRTRSWSEFGLAHATNATLAFCLFGLFVRSOPCYNRARGTLFAMLAETITWSEVP 780
DB 721 HCRTRSWSEFGLAHATNATLAFCLFGLFVRSOPCYNRARGTLFAMLAETITWSEVP 780
QY 781 LLANVOVYLPAVOMGALLCYGILAAEHLPRCYLLMRQPLNTPPEFLGGGPGDAQO 840
DB 781 LLANVOVYLPAVOMGALLCYGILAAEHLPRCYLLMRQPLNTPPEFLGGGPGDAQO 840
QY 841 NDNGTNGNCKHE 852
DB 841 NDNGTNGNCKHE 852

```

RESULT 4

```

US-10-261-482-2
; Sequence 2, Application US/10261482
; Publication No. US20030036089A1
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al
; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
; TITLE OF INVENTION: RECEPTORS, NUCLEIC ACID MOLECULES ENCODING HUMAN GPCR
; FILE REFERENCE: CL000869CON
; CURRENT APPLICATION NUMBER: US/10/261,482
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: 09/684,393
; PRIOR FILING DATE: 2000-10-10
; PRIOR APPLICATION NUMBER: 60/172,600
; PRIOR FILING DATE: 1999-12-20
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 852
; TYPE: PRT
; ORGANISM: HUMAN
US-10-261-482-2

```

```

Query Match      99.7%; Score 4512; DB 9; Length 852;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 851; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 MCGPAVLGSLTMAALLHPGTGAPCLISQQLRMKGDYVIGLFLPGAEEAGLSRTRPSSP 60
DB 1 MCGPAVLGSLTMAALLHPGTGAPCLISQQLRMKGDYVIGLFLPGAEEAGLSRTRPSSP 60
QY 61 VCTRESSNGLTMAALMKAAVEEINNKSDLLPGIRLGIDYDFDCSEPVVAMKPSLMFLAKA 120
DB 61 VCTRESSNGLTMAALMKAAVEEINNKSDLLPGIRLGIDYDFDCSEPVVAMKPSLMFLAKA 120
QY 121 GSRDIAAYCNYOYOPRYAVIGPHSSELAAMYTGKFFSEFLMPQVSYGSMELLSARETF 180
DB 121 GSRDIAAYCNYOYOPRYAVIGPHSSELAAMYTGKFFSEFLMPQVSYGSMELLSARETF 180
QY 181 PSFFRTVPDRVQVTAALAEELQEFQWNVVAAALGSDDEYGRGISTIFSALAAAGCICIAHE 240
DB 181 PSFFRTVPDRVQVTAALAEELQEFQWNVVAAALGSDDEYGRGISTIFSALAAAGCICIAHE 240
QY 241 GLVPLPRADDSRLGKYQVDLHOVNOSSVOYVLLFASVHAHAHLFNTSISRSPKXWVAS 300
DB 241 GLVPLPRADDSRLGKYQVDLHOVNOSSVOYVLLFASVHAHAHLFNTSISRSPKXWVAS 300
QY 301 EAMLTSDLVMGLPMAOMGTVLGFLORGAQLHEFPQVYVTHLALATDPAFCALGEREG 360
DB 301 EAMLTSDLVMGLPMAOMGTVLGFLORGAQLHEFPQVYVTHLALATDPAFCALGEREG 360
QY 361 LEEDVVGQRCPQDCITLTONVSAGLNHQTFSVYAAVYSVAQALNHTLQCNASGCPADP 420
DB 361 LEEDVVGQRCPQDCITLTONVSAGLNHQTFSVYAAVYSVAQALNHTLQCNASGCPADP 420
QY 421 VAPWOLLEMMYLTFRVHGGLPLRFDSGVDMEYDKLVMWOGSVPRLDHVRFNGLRT 480
DB 421 VAPWOLLEMMYLTFRVHGGLPLRFDSGVDMEYDKLVMWOGSVPRLDHVRFNGLRT 480

```

QY	481	ERLIRNHTSNOKPVCSCPOCGGVRRKPGHSCCYCVCDEAGSYRONDDIACF	540
Db	481	ERLIRNHTSNOKPVCSCPOCGGVRRKPGHSCCYCVCDEAGSYRONDDIACF	540
QY	541	CGDEWSPENSTRCFRRRSRFLWGEPAVLLLLLSIALGLVLAALGLFVHHRSDPLYO	600
Db	541	CGDEWSPENSTRCFRRRSRFLWGEPAVLLLLLSIALGLVLAALGLFVHHRSDPLYO	600
QY	601	ASGGPLACFGLVLCGLVCLSVLFLPGQPSPARCIAQOPLSHDLPTGCLSTFLQAAEIFY	660
Db	601	ASGGPLACFGLVLCGLVCLSVLFLPGQPSPARCIAQOPLSHDLPTGCLSTFLQAAEIFY	660
QY	661	ESELPLSMADLSGCLGPMAMLVLLIAMVEVALCTWYLVAPPEVYTDWMLPTEALY	720
Db	661	ESELPLSMADLSGCLGPMAMLVLLIAMVEVALCTWYLVAPPEVYTDWMLPTEALY	720
QY	721	HCTRSMWSEFGLAATNATLAFLOCTGLTVRSQPGCYNRARGITFEMLAIFYFTWVSFVP	780
Db	721	HCTRSMWSEFGLAATNATLAFLOCTGLTVRSQPGCYNRARGITFEMLAIFYFTWVSFVP	780
QY	781	LLANVOYVLRPAVOMGAILLCVGLTAAFIHLPRCYLLMKRPGILNTPPEFLYGGGPGDAQO	840
Db	781	LLANVOYVLRPAVOMGAILLCVGLTAAFIHLPRCYLLMKRPGILNTPPEFLYGGGPGDAQO	840
QY	841	NDGNTNGOGKHE	852
Db	841	NDGNTNGOGKHE	852

```

RESULT 5
US-10-282-837-14
: Sequence 14, Application US/10282837
: Publication No. US20030082738A1
: GENERAL INFORMATION:
: APPLICANT: Millennium Pharmaceuticals, Inc.
: TITLE OF INVENTION: 1983, 52881, 2398, 45449, 50289, AND
: TITLE OF INVENTION: 52872, NOVEL G PROTEIN-COUPLED RECEPTORS AND USES THEREFOR
: FILE REFERENCE: 10448-020001
: CURRENT APPLICATION NUMBER: US/10/282,837
: CURRENT FILING DATE: 2002-10-29
: PRIOR APPLICATION NUMBER: US/09/796,338
: PRIOR FILING DATE: 2001-02-28
: PRIOR APPLICATION NUMBER: US 60/186,059
: PRIOR FILING DATE: 2000-02-29
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 14
: LENGTH: 852
: TYPE: PRF
: ORGANISM: Homo sapiens
: US-10-282-837-14

```

Query Match	99.7%	Score 4512	DB 9	Length 852
Best Local Similarity	99.9%	Pred No. 0		
Matches	851	Conservative 0	Mismatches 1	Indels 0
			Gaps	0
QY	1	MGPAYVLGLSLMALHPEGTGAPLCISQOLRMKGVDYVGLFPLFGEAEAGLRSTRPSSP	60	
Db	1	MGPVYVLGLSLMALHPEGTGAPLCISQOLRMKGVDYVGLFPLFGEAEAGLRSTRPSSP	60	
QY	61	VCTRSSNGILMALMKKAAVEEINNKSDLPGLRIGYDLFDPCSEPVYAMKRSIMFLIAA	120	
Db	61	VCTRSSNGILMALMKKAAVEEINNKSDLPGLRIGYDLFDPCSEPVYAMKRSIMFLIAA	120	
QY	121	GSRDIAAYCANTQYOYPRVLAVTGPSSSELAMVTGKFFSEFLMPQVSYGASMELLSARETF	180	
Db	121	GSRDIAAYCANTQYOYPRVLAVTGPSSSELAMVTGKFFSEFLMPQVSYGASMELLSARETF	180	
QY	181	PSFFRTVSDRQVLAAAEILQEFGMNWWAALGSDDEYRGQSLSTFSALAAARGICIAHE	240	
Db	181	PSFFRTVSDRQVLAAAEILQEFGMNWWAALGSDDEYRGQSLSTFSALAAARGICIAHE	240	
QY	241	GLVPLPRADDSTRGKYQDVVLHOVNOSSVOYVLLFASVHAHAALFNYSISSRLSPKWVAS	300	

Db	241	GLVPLPRADSRGLKQYDVLHQVWSSQYVLLFASVHAANALFNYSISSRLSPKVVAS	3000
Qy	301	EAMLTSDLVMLPGMAQMGUYLGLFQGAOLHERPQYUKYHIALATPAPCSALGEREGG	3600
Db	301	EAMLTSDLVMLPGMAQMGUYLGLFQGAOLHERPQYUKYHIALADPACSAALGEREGG	3600
Qy	361	LEEDVQGRQCCQCCITLQNVSAALNHQFTSVYAAYVSAQALNHTLQCNASGCPKODP	4200
Db	361	LEEDVQGRQCCQCCITLQNVSAALNHQFTSVYAAYVSAQALNHTLQCNASGCPKODP	4200
Qy	421	VKPMQLENNMTNLFHFVHGCLPLRPDSSGNVMEYDLKLMWQSGVPRLHDVGRNSLRT	4800
Db	421	VKPMQLENNMTNLFHFVHGCLPLRPDSSGNVMEYDLKLMWQSGVPRLHDVGRNSLRT	4800
Qy	481	ERLKRMTSDNQRPVSCSGRQCCGGVRRVKGHSCCYCVCCEASYSRQNPDDIACFE	5400
Db	481	ERLKRMTSDNQRPVSCSGRQCCGGVRRVKGHSCCYCVCCEASYSRQNPDDIACFE	5400
Qy	541	CGODEMSPERSTRCQFRRRSRRLANGEBRAVLLLLLSLALGVTLAALGLRVHHRDSTLVQ	6000
Db	541	CGODEMSPERSTRCQFRRRSRRLANGEBRAVLLLLLSLALGVTLAALGLRVHHRDSTLVQ	6000
Qy	601	ASGGFLACFGVLCGLVCLSVLPLPGQSPSPARCLAQOPLSHLPJTGCLSTFLQAAEFV	6600
Db	601	ASGGFLACFGVLCGLVCLSVLPLPGQSPSPARCLAQOPLSHLPJTGCLSTFLQAAEFV	6600
Qy	661	ESEPLSMADRLSGCLKGPAMLVYLLAMLYEVALCTWYLAAPPEVYVTHMHLPTALV	7200
Db	661	ESEPLSMADRLSGCLKGPAMLVYLLAMLYEVALCTWYLAAPPEVYVTHMHLPTALV	7200
Qy	721	HCRTRSWSPFGLAHATNATLALFCFLGTFLVRSQPGCYNBARGLTFMALAYFTWVSFP	7800
Db	721	HCRTRSWSPFGLAHATNATLALFCFLGTFLVRSQPGYNRARGTFMALAYFTWVSFP	7800
Qy	781	LIANVQVLRPAVOMGALLICVLGTLAFLPRCTYLLMRQGLNTPEFTLGGGPGDAQO	8400
Db	781	LIANVQVLRPAVOMGALLICVLGTLAFLPRCTYLLMRQGLNTPEFTLGGGPGDAQO	8400
Qy	841	NDGNTNGNGKHE 852	
Db	841	NDGNTNGNGKHE 852	

```

RESULT 6
US-09-796-338A-14
: Sequence 14, Application US/09796338A
: Patent No. US20020061522A1
: GENERAL INFORMATION:
: APPLICANT: Millennium Pharmaceuticals, Inc.
: TITLE OF INVENTION: 1983, 52881, 2398, 45449, 50289, AND
: TITLE OF INVENTION: 52872, NOVEL G PROTEIN-COUPLED RECEPTORS AND USES THEREFOR
: FILE REFERENCE: 10448-020001
: CURRENT APPLICATION NUMBER: US/09/796,338A
: CURRENT FILING DATE: 2001-02-28
: PRIOR APPLICATION NUMBER: 2000-02-29
: PRIOR FILING DATE: 2000-02-29
: NUMBER OF SEQ ID NOS: 26
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 14
: LENGTH: 852
: TYPE: PR1
: ORGANISM: Homo sapiens
: US-09-796-338A-14

```

Query Match	Similarity	Score	DB	Length	852;
Best Local	Similarity	99.9%	Pred.	No. 0;	
Matches	851;	Conservative	0;	Mismatches	1;
			Indels	0;	Gaps
QY	1	MIGPVLIGLSIMALLHRTGAPLCLCSQQLRMKGDIYVIGLPIPLGAEEAFGLRSPRRSSP	60		
Db	1	MIGPVLIGLSIMALLHRTGAPLCLCSQQLRMKGDIYVIGLPIPLGAEEAFGLRSPRRSSP	60		

QY 61 VCTRESSNGLLMALAMKAABEINNKSDLLPGLRGYDLFTDCSEPPVAMKPSLMFLAKA 120
D61 VCTRESSNGLLMALAMKAABEINNKSDLLPGLRGYDLFTDCSEPPVAMKPSLMFLAKA 120
QY 121 GSRDIAACNTQYQOPRYLAVIIGHSSSELAWTGKFFSFPLMPQVSYGASMEILLSARETF 180
D121 GSRDIAACNTQYQOPRYLAVIIGHSSSELAWTGKFFSFPLMPQVSYGASMEILLSARETF 180
QY 181 PSFRTVSDRYOLTAALAEILQEFEGMNVVALGSDDEYRGGLSTFSALAAAGICIAHE 240
D181 PSFRTVSDRYOLTAALAEILQEFEGMNVVALGSDDEYRGGLSTFSALAAAGICIAHE 240
QY 241 GLVPLPRADDSRLGKVQDYLHQVNOSSVQVYLLFASVAHAALFNYSISSRLSPRWVWAS 300
D241 GLVPLPRADDSRLGKVQDYLHQVNOSSVQVYLLFASVAHAALFNYSISSRLSPRWVWAS 300
QY 301 EAMLTSDLVNGLPGMAQNGTVLGFQRGALHEFPQYKTHALATDPAFCALGERBQG 360
D301 EAMLTSDLVNGLPGMAQNGTVLGFQRGALHEFPQYKTHALATDPAFCALGERBQG 360
QY 361 LEEVDVGQRCPQCCITLQNVASAGLNHHQTFSVYAAYVSAQALHNTLQCNASGCCPADP 420
D361 LEEVDVGQRCPQCCITLQNVASAGLNHHQTFSVYAAYVSAQALHNTLQCNASGCCPADP 420
QY 421 VKPQOLLENMTNLTFFHVAGLPLRFDSGSGNVDMEXDLKLMWOGSVPRLDHVGRENGSLRT 480
D421 VKPQOLLENMTNLTFFHVAGLPLRFDSGSGNVDMEXDLKLMWOGSVPRLDHVGRENGSLRT 480
QY 481 ERLKIRWHTSDNOKPVSRCSRQCEGQYRVKGFHSCCYDCVDCAGSYRONPDDIACTF 540
D481 ERLKIRWHTSDNOKPVSRCSRQCEGQYRVKGFHSCCYDCVDCAGSYRONPDDIACTF 540
QY 541 CGODEMSPERSTRCFRRRSRFLANGEPAYVLLLLSTALGLVLAALGLFYHHRDSPLVQ 600
D541 CGODEMSPERSTRCFRRRSRFLANGEPAYVLLLLSTALGLVLAALGLFYHHRDSPLVQ 600
QY 601 ASGGPLACFGVLGVLCLVSLVLPFGQSPARCLAQOPLSHLPLTGLCSTLFLQAAELFV 660
D601 ASGGPLACFGVLGVLCLVSLVLPFGQSPARCLAQOPLSHLPLTGLCSTLFLQAAELFV 660
QY 661 ESELPLSWADRLSGCLRGPAMLVLLAMLEVVALCTYVLAFFPEVYTDHMLPTEALV 720
D661 ESELPLSWADRLSGCLRGPAMLVLLAMLEVVALCTYVLAFFPEVYTDHMLPTEALV 720
QY 721 HCRTRSWVSGFLAATNATLAFCLFGLTFLVRSQPGCYNRARGLTFEALAFITWVSVVP 780
D721 HCRTRSWVSGFLAATNATLAFCLFGLTFLVRSQPGCYNRARGLTFEALAFITWVSVVP 780
QY 781 LIANVQVVLPRAYOMGALLLCVLGILAAFIHLPCLYLMRQPGMLTPEFFLGSGGPDAGQ 840
D781 LIANVQVVLPRAYOMGALLLCVLGILAAFIHLPCLYLMRQPGMLTPEFFLGSGGPDAGQ 840
QY 841 NDGNTNGNGKHE 852
D841 NDGNTNGNGKHE 852

RESULT 7
US-09-799-629-4

; Sequence 4, Application US/09799629
; Publication No. US20030008344A1
; GENERAL INFORMATION:
; APPLICANT: ADLER, JON ELLIOT
; APPLICANT: ZOZULYA, SERGEY
; APPLICANT: LI, XIADONG
; APPLICANT: O'CONNELL, SHAWN
; APPLICANT: STASZEMSKI, LENA
; TITLE OF INVENTION: T1R TASTE RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: 078003/027870/RXT
; CURRENT APPLICATION NUMBER: US/09/799,629
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: 60/187,546
; PRIOR FILING DATE: 2000-03-07

; PRIOR APPLICATION NUMBER: 60/195,536
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: 60/209,840
; PRIOR FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: 60/214,213
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 60/226,448
; PRIOR FILING DATE: 2000-08-17
; PRIOR APPLICATION NUMBER: 60/259,227
; PRIOR FILING DATE: 2001-01-03
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 850
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-799-629-4
Query Match 99.5%; Score 4500; DB 9; Length 850;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 849; Conservative 0; Mismatches 1; Indels 2; Gaps 1;
QY 1 MLGPAYLGLSLMALHPGTGAPLCLSQGLRMKGDVYLGLEPLGAEBAEAGLRSTRPESP 60
D1 MLGPAYLGLSLMALHPGTGAPLCLSQGLRMKGDVYLGLEPLGAEBAEAGLRSTRPESP 60
QY 61 VCTRESSNGLLMALAMKAABEINNKSDLLPGLRGYDLFTDCSEPPVAMKPSLMFLAKA 120
D61 VCTRESSNGLLMALAMKAABEINNKSDLLPGLRGYDLFTDCSEPPVAMKPSLMFLAKA 120
QY 121 GSRDIAACNTQYQOPRYLAVIIGHSSSELAWTGKFFSFPLMPQVSYGASMEILLSARETF 180
D121 GSRDIAACNTQYQOPRYLAVIIGHSSSELAWTGKFFSFPLMPH - YASMEILLSARETF 178
QY 181 PSFRTVSDRYOLTAALAEILQEFEGMNVVALGSDDEYRGGLSTFSALAAAGICIAHE 240
D181 PSFRTVSDRYOLTAALAEILQEFEGMNVVALGSDDEYRGGLSTFSALAAAGICIAHE 238
QY 241 GLVPLPRADDSRLGKVQDYLHQVNOSSVQVYLLFASVAHAALFNYSISSRLSPRWVWAS 300
D241 GLVPLPRADDSRLGKVQDYLHQVNOSSVQVYLLFASVAHAALFNYSISSRLSPRWVWAS 298
QY 299 EAMLTSDLVNGLPGMAQNGTVLGFQRGALHEFPQYKTHALATDPAFCALGERBQG 358
D299 EAMLTSDLVNGLPGMAQNGTVLGFQRGALHEFPQYKTHALATDPAFCALGERBQG 358
QY 359 LEEVDVGQRCPQCCITLQNVASAGLNHHQTFSVYAAYVSAQALHNTLQCNASGCCPADP 418
D359 LEEVDVGQRCPQCCITLQNVASAGLNHHQTFSVYAAYVSAQALHNTLQCNASGCCPADP 418
QY 421 VKPQOLLENMTNLTFFHVAGLPLRFDSGSGNVDMEXDLKLMWOGSVPRLDHVGRENGSLRT 480
D421 VKPQOLLENMTNLTFFHVAGLPLRFDSGSGNVDMEXDLKLMWOGSVPRLDHVGRENGSLRT 478
QY 479 ERLKIRWHTSDNOKPVSRCSRQCEGQYRVKGFHSCCYDCVDCAGSYRONPDDIACTF 538
D479 ERLKIRWHTSDNOKPVSRCSRQCEGQYRVKGFHSCCYDCVDCAGSYRONPDDIACTF 538
QY 541 CGODEMSPERSTRCFRRRSRFLANGEPAYVLLLLSTALGLVLAALGLFYHHRDSPLVQ 600
D541 CGODEMSPERSTRCFRRRSRFLANGEPAYVLLLLSTALGLVLAALGLFYHHRDSPLVQ 600
QY 601 ASGGPLACFGVLGVLCLVSLVLPFGQSPARCLAQOPLSHLPLTGLCSTLFLQAAELFV 660
D601 ASGGPLACFGVLGVLCLVSLVLPFGQSPARCLAQOPLSHLPLTGLCSTLFLQAAELFV 658
QY 661 ESELPLSWADRLSGCLRGPAMLVLLAMLEVVALCTYVLAFFPEVYTDHMLPTEALV 720
D661 ESELPLSWADRLSGCLRGPAMLVLLAMLEVVALCTYVLAFFPEVYTDHMLPTEALV 718
QY 721 HCRTRSWVSGFLAATNATLAFCLFGLTFLVRSQPGCYNRARGLTFEALAFITWVSVVP 780
D721 HCRTRSWVSGFLAATNATLAFCLFGLTFLVRSQPGCYNRARGLTFEALAFITWVSVVP 778

QY 781 L1ANVOVLPRAVOMGALLICVIGI1IAAFHLPRCYLLMRPGILMTPEFLGGGPGDAQO 840
 Db 779 L1ANVOVLPRAVOMGALLICVIGI1IAAFHLPRCYLLMRPGILMTPEFLGGGPGDAQO 838
 QY 841 NDGNTGNQGRHE 852
 Db 839 NDGNTGNQGRHE 850

RESULT 8

US-09-799-629-14
 ; Sequence 14, Application US/09799629
 ; Publication No. US20030008344A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ADLER, JON ELLIOT
 ; APPLICANT: ZOZULYA, SERGEY
 ; APPLICANT: LI, XIADONG
 ; APPLICANT: O'CONNELL, SHAWN
 ; APPLICANT: STASZEMSKI, LENA
 ; TITLE OF INVENTION: T1R TASTE RECEPTORS AND GENES ENCODING SAME
 ; FILE REFERENCE: 078003/0277870/RXT
 ; CURRENT APPLICATION NUMBER: US/09/799, 629
 ; PRIOR FILING DATE: 2001-03-07
 ; PRIOR APPLICATION NUMBER: 60/187,546
 ; PRIOR FILING DATE: 2000-03-07
 ; PRIOR APPLICATION NUMBER: 60/195,536
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: 60/209,840
 ; PRIOR FILING DATE: 2000-06-06
 ; PRIOR APPLICATION NUMBER: 60/214,213
 ; PRIOR FILING DATE: 2000-06-23
 ; PRIOR APPLICATION NUMBER: 60/226,448
 ; PRIOR FILING DATE: 2000-08-17
 ; PRIOR APPLICATION NUMBER: 60/259,227
 ; PRIOR FILING DATE: 2001-01-03
 ; NUMBER OF SEQ ID NOS: 20
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 14
 ; LENGTH: 858
 ; TYPE: PRT
 ; ORGANISM: Rattus sp.
 US-09-799-629-14

Query Match 71.9%; Score 3255; DB 9; Length 858;
 Best Local Similarity 72.9%; Pred. No. 2.8e-254;
 Matches 623; Conservative 74; Mismatches 144; Indels 14; Gaps 4;

QY 1 MLCRAVIGISIMALLHGTGAPCLISQOLRMKGDYVIGLFLPGAEEAGLRSTRSPSP 60
 Db 1 MGLAIIIGLSIAFLLEIGMSISLISQOFKAGDYIIIGLFLPGAEEAGLRSTRSPSP 60
 QY 61 VCTFRSSNGLLMALAMKAAVEINNKSDLPGLRGYDLFTPTCEPVPYAMPSLMTAKA 120
 Db 61 LCTFRSPGLFLAMAMKAAVEINNKSDLPGLRGYDLFTPTCEPVPYAMPSLMTAKA 120
 QY 121 GSRDIAAYCANTOYQPRVLAVIIGHSESLANVTGKFSFFLMPQVSYGASMSLARETF 180
 Db 121 GSOSIAAYCANTOYQPRVLAVIIGHSESLANVTGKFSFFLMPQVSYGASMSLARETF 180
 QY 181 PSFRTPVSDRVOYLTAAAEILLEGCGWVVALGSDDEYRGGLSIFSLAARICIAHE 240
 Db 181 PSFRTPVSDRVOYLTAAAEILLEGCGWVVALGSDDEYRGGLSIFSLAARICIAHE 240
 QY 241 GLVPLPRADBSRLGKODVYLHOVQSSVOYVLLPASYHAALHNSYSSLSKRVWVAS 300
 Db 241 GLVQOHTSGOGLKQVVDVLRQVQSKVQYVVLPAASARAYSLSTSLHDLSPRVWVAS 300
 QY 301 EAWLTSDLVMLPGMAQMTVIGLQGAOLHEPPQYKTHLALATDPACFALGERBOG 360
 Db 301 ESWLTSDLVMLPNIARVGLVIGLQGALLPERSHYETRLALADPTFCASL-KAELD 359
 QY 361 LEEVYQRCRCQCCITLQNVASAGLN-----HQTFSYAAVYSVAQALHNTLQCNV 412

Db 360 LEEVYQRCRCQCCITLQNVASAGLN-----HQTFSYAAVYSVAQALHNTLQCNV 419
 QY 413 SCGPADQVAPRMOLLENNYMLTFHVGLPLRFDSNGVNDYEDLKLWMOGSPRLHDVG 472
 Db 420 SHCHTSEVPQWOLENNYMKSFARADLTQFDKAGSYVDHEYDLKMWKSPPTPVHTVG 479
 QY 473 RFNGSLTERIKIRMTSDNOKPVSRCSROCBQVRRVKGFSHCYDVCENGSTRON 532
 Db 480 TFNGTLOHRSKMYW--PGNOVPVSCSROCKDQVRRVKGFSHCYDVCENGSTRON 537
 QY 533 PDDIACFFCGODEWSPERSRRCFRRRSRFLAMGPATVLLIILSTALGLVLAAGLFVH 592
 Db 538 PDDIACFFCGODEWSPERSRRCFRRRSRFLAMGPATVLLIILSTALGLVLAAGLFVH 597
 QY 593 HRDSPLOVAGSGPLACGLVLCGLVSVLLFPGQSPARCIAQOPLSHPLTGCLSTLF 652
 Db 598 YWDSPLVQASGSLFCGLICLIGLFCVLLFPRPSPASACIAQOPLSHPLTGCLSTLF 657
 QY 653 LQAAEFVESELPLSMADRLSGCLRGWMLVLLAMLEVALCTWYLVAPPEVYDWH 712
 Db 658 LQAAEFVESELPLSMADRLSGCLRGWMLVLLAMLEVALCTWYLVAPPEVYDWH 717
 QY 713 MLPTFAVHCRTBSWVSEGLAHATNATLACPLGTIVYSQRCYRARGITPAMLATF 772
 Db 718 VLPTEVLEHCRMRWSVSLGVHTINAVLAFLCFLGTIVYSQRCYRARGITPAMLATF 777
 QY 773 ITWVSFPLLANVOVLPRAVOMGALLICVIGI1IAAFHLPRCYLLMRPGILMTPEFLGG 832
 Db 778 ITWVSFPLLANVOVLPRAVOMGALLICVIGI1IAAFHLPRCYLLMRPGILMTPEFLGG 837
 QY 833 GPGDAGQDNDGNTGN 847
 Db 838 SPKEA---SDGNSGS 849

RESULT 9

US-09-927-315-25
 ; Sequence 25, Application US/09927315
 ; Publication No. US20030040045A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Zuker, Charles S.
 ; APPLICANT: Ryba, Nicholas J.P.
 ; APPLICANT: Nelson, Greg
 ; APPLICANT: Hoon, Mark A.
 ; APPLICANT: Chandrasekar, Jayaram
 ; APPLICANT: Zhang, Yifeng
 ; APPLICANT: The Regents of the University of California
 ; APPLICANT: The Government of the United States of America
 ; APPLICANT: as represented by the Secretary of the
 ; TITLE OF INVENTION: Mammalian Sweet Taste Receptors
 ; FILE REFERENCE: 02307E-120110US
 ; CURRENT APPLICATION NUMBER: US/09/927,315
 ; PRIOR FILING DATE: 2001-08-10
 ; PRIOR APPLICATION NUMBER: US 60/302,898
 ; PRIOR FILING DATE: 2001-07-03
 ; NUMBER OF SEQ ID NOS: 25
 ; SOFTWARE: Patent In Ver. 2.1
 ; SEQ ID NO 25
 ; LENGTH: 858
 ; TYPE: PRT
 ; ORGANISM: Rattus sp.
 ; FEATURE:
 ; OTHER INFORMATION: rat T1R3 sweet taste receptor
 US-09-927-315-25

Query Match 71.9%; Score 3255; DB 9; Length 858;
 Best Local Similarity 72.9%; Pred. No. 2.8e-254;
 Matches 623; Conservative 74; Mismatches 144; Indels 14; Gaps 4;
 QY 1 MLCRAVIGISIMALLHGTGAPCLISQOLRMKGDYVIGLFLPGAEEAGLRSTRSPSP 60
 Db 1 MGLAIIIGLSIAFLLEIGMSISLISQOFKAGDYIIIGLFLPGAEEAGLRSTRSPSP 60

```

Db 1 MGLAIIIGLSIAFLAFLGSSSLCSQOFKAGDYLIGLFLPGTTEATLNRQPNIGI 60
QY 61 VCTRSSNGLMALAMKMAVEEINKSDLLPGLRLGYDLFPTCSEPVYAMKSLMFLA 120
Db 61 ICTRSSPLGLFLAMKMAVEEINKSALLPGLRLGYDLFPTCSEPVYAMKSLMFLA 120
QY 121 GSRDIAAYCNTYOYOPRYLAIVIGPHSESLAATGKFFSFPLMPQVYSASMRISDRETF 180
Db 121 GSQSIATAYCNTYOYOPRYLAIVIGPHSESLAATGKFFSFPLMPQVYSASMRISDRETF 180
QY 181 PSFRTVPDSRVOLTAALAELOEFGNNVVAALGSDEYGRGSLIFSALAAAGCIIAHE 240
Db 181 PSFRTVPDSRVOLTAALAELOEFGNNVVAALGSDEYGRGSLIFSALAAAGCIIAHE 240
QY 241 GLVPLPRADDSRLGKQVODLHOVNOSQVYVLLFASVHAHAHLFNYSISSRLSPKRWAS 300
Db 241 GLVPLPRADDSRLGKQVODLHOVNOSQVYVLLFASVHAHAHLFNYSISSRLSPKRWAS 300
QY 301 EAMLTSDLVYMGJLPGMAOMGTVLGFLQRGALHEFPQYVYTHALATDPAFCALGEREG 360
Db 301 EAMLTSDLVYMGJLPGMAOMGTVLGFLQRGALHEFPQYVYTHALATDPAFCALGEREG 360
QY 361 LEEDVVGORCPQCCITLQNVASGLN-----HHQFTSYAAYVSAQAHLNLTQCN 412
Db 361 LEEDVVGORCPQCCITLQNVASGLN-----HHQFTSYAAYVSAQAHLNLTQCN 412
QY 413 SGCPADDPVKPQMLENNYNLTFEHVGLPLRDSNGVNDMEVDLKMVWQSVPLRHVG 472
Db 413 SGCPADDPVKPQMLENNYNLTFEHVGLPLRDSNGVNDMEVDLKMVWQSVPLRHVG 472
QY 473 RENGLRTELRKIRWHTSDNOKPVSRCSROCEGQVRYKGFHSCCYDCVDCAGSYRON 532
Db 473 RENGLRTELRKIRWHTSDNOKPVSRCSROCEGQVRYKGFHSCCYDCVDCAGSYRON 532
QY 532 PDDICTFCGQDQWSPERSTCTCLPRPKFLANGEPVAVLSLLLLGLVGLTAAALGLFVH 597
Db 532 PDDICTFCGQDQWSPERSTCTCLPRPKFLANGEPVAVLSLLLLGLVGLTAAALGLFVH 597
QY 593 HRDSPVQASGGPLACFLGVLGVLCLSVLLEPGOPSPARCIAOQPLSLPLTGLSTLF 652
Db 593 HRDSPVQASGGPLACFLGVLGVLCLSVLLEPGOPSPARCIAOQPLSLPLTGLSTLF 652
QY 653 LOAAEIFESELPLSMANMLCSYLRGPAMLVLLATLVEALCAMYLMAFPPEVYTDNQ 717
Db 653 LOAAEIFESELPLSMANMLCSYLRGPAMLVLLATLVEALCAMYLMAFPPEVYTDNQ 717
QY 713 MLPTALVHCRTSRVSWFGLAHATNATLAFCLGTFELVRSOPGCYNRARGITFAMLAYF 772
Db 713 MLPTALVHCRTSRVSWFGLAHATNATLAFCLGTFELVRSOPGCYNRARGITFAMLAYF 772
QY 773 ITWVSFVPLANVOYVLRPAVOMGALLCIVGLIILAFHLPRCYLMLRQGLMTPPEFLG 832
Db 773 ITWVSFVPLANVOYVLRPAVOMGALLCIVGLIILAFHLPRCYLMLRQGLMTPPEFLG 832
QY 833 GPDAOGQNDGNTGN 847
Db 833 SPKRA---SDNGSGS 849

```

```

RESULT 10
US-10-035-045-14
; Sequence 14, Application US/10035045
; Publication No. US20030054448A1
; GENERAL INFORMATION:
; APPLICANT: ADLER, JON ELLIOT
; APPLICANT: LI, XIAODONG
; APPLICANT: STASZEWSKI, LENA
; APPLICANT: O'CONNELL, SHAWN
; APPLICANT: ZOZULYA, SERGEY
; TITLE OF INVENTION: TIR TASTE RECEPTORS AND GENES ENCODING SAME
; FILE REFERENCE: 078003-0280681
; CURRENT APPLICATION NUMBER: US/10/035, 045
; CURRENT FILING DATE: 2002-01-03

```

```

; PRIOR APPLICATION NUMBER: 60/259, 227
; PRIOR FILING DATE: 2001-01-03
; PRIOR APPLICATION NUMBER: 60/284, 547
; PRIOR FILING DATE: 2001-04-19
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 858
; TYPE: PRT
; ORGANISM: Rattus sp.
; US-10-035-045-14

Query Match          71.9%  Score 3255;  DB 9;  Length 858;
Best Local Similarity 72.9%  Pred. No. 2,8e-254;
Matches 623; Conservative 74; Mismatches 144; Indels 14; Gaps 4;

QY 1 MGLPVLIGLSIAALLHPTGAPLCLSQOIRKMGDYLIGLFLPGEAEEGLNSRTPSSP 60
Db 1 MGLAIIIGLSIAFLAFLGSSSLCSQOFKAGDYLIGLFLPGTTEATLNRQPNIGI 60
QY 61 VCTRSSNGLMALAMKMAVEEINKSDLLPGLRLGYDLFPTCSEPVYAMKSLMFLA 120
Db 61 ICTRSSPLGLFLAMKMAVEEINKSALLPGLRLGYDLFPTCSEPVYAMKSLMFLA 120
QY 121 GSRDIAAYCNTYOYOPRYLAIVIGPHSESLAATGKFFSFPLMPQVYSASMRISDRETF 180
Db 121 GSQSIATAYCNTYOYOPRYLAIVIGPHSESLAATGKFFSFPLMPQVYSASMRISDRETF 180
QY 181 PSFRTVPDSRVOLTAALAELOEFGNNVVAALGSDEYGRGSLIFSALAAAGCIIAHE 240
Db 181 PSFRTVPDSRVOLTAALAELOEFGNNVVAALGSDEYGRGSLIFSALAAAGCIIAHE 240
QY 241 GLVPLPRADDSRLGKQVODLHOVNOSQVYVLLFASVHAHAHLFNYSISSRLSPKRWAS 300
Db 241 GLVPLPRADDSRLGKQVODLHOVNOSQVYVLLFASVHAHAHLFNYSISSRLSPKRWAS 300
QY 301 EAMLTSDLVYMGJLPGMAOMGTVLGFLQRGALHEFPQYVYTHALATDPAFCALGEREG 360
Db 301 EAMLTSDLVYMGJLPGMAOMGTVLGFLQRGALHEFPQYVYTHALATDPAFCALGEREG 360
QY 361 LEEDVVGORCPQCCITLQNVASGLN-----HHQFTSYAAYVSAQAHLNLTQCN 412
Db 361 LEEDVVGORCPQCCITLQNVASGLN-----HHQFTSYAAYVSAQAHLNLTQCN 412
QY 413 SGCPADDPVKPQMLENNYNLTFEHVGLPLRDSNGVNDMEVDLKMVWQSVPLRHVG 472
Db 413 SGCPADDPVKPQMLENNYNLTFEHVGLPLRDSNGVNDMEVDLKMVWQSVPLRHVG 472
QY 473 RENGLRTELRKIRWHTSDNOKPVSRCSROCEGQVRYKGFHSCCYDCVDCAGSYRON 532
Db 473 RENGLRTELRKIRWHTSDNOKPVSRCSROCEGQVRYKGFHSCCYDCVDCAGSYRON 532
QY 532 PDDICTFCGQDQWSPERSTCTCLPRPKFLANGEPVAVLSLLLLGLVGLTAAALGLFVH 597
Db 532 PDDICTFCGQDQWSPERSTCTCLPRPKFLANGEPVAVLSLLLLGLVGLTAAALGLFVH 597
QY 593 HRDSPVQASGGPLACFLGVLGVLCLSVLLEPGOPSPARCIAOQPLSLPLTGLSTLF 652
Db 593 HRDSPVQASGGPLACFLGVLGVLCLSVLLEPGOPSPARCIAOQPLSLPLTGLSTLF 652
QY 653 LOAAEIFESELPLSMANMLCSYLRGPAMLVLLATLVEALCAMYLMAFPPEVYTDNQ 717
Db 653 LOAAEIFESELPLSMANMLCSYLRGPAMLVLLATLVEALCAMYLMAFPPEVYTDNQ 717
QY 713 MLPTALVHCRTSRVSWFGLAHATNATLAFCLGTFELVRSOPGCYNRARGITFAMLAYF 772
Db 713 MLPTALVHCRTSRVSWFGLAHATNATLAFCLGTFELVRSOPGCYNRARGITFAMLAYF 772
QY 773 ITWVSFVPLANVOYVLRPAVOMGALLCIVGLIILAFHLPRCYLMLRQGLMTPPEFLG 832
Db 773 ITWVSFVPLANVOYVLRPAVOMGALLCIVGLIILAFHLPRCYLMLRQGLMTPPEFLG 832
QY 833 GPDAOGQNDGNTGN 847

```


RESULT 14
 US-10-261-482-4
 ; Sequence 4, Application US/10261482
 ; Publication No. US20030036089A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WEI, Ming-Hui et al
 ; TITLE OF INVENTION: ISOLATED HUMAN G-PROTEIN COUPLED
 ; TITLE OF INVENTION: RECEPTORS, NOCLEIC ACID MOLECULES ENCODING HUMAN GPCR
 ; FILE REFERENCE: CLO00869CON
 ; CURRENT APPLICATION NUMBER: US/10/261,482
 ; PRIOR APPLICATION NUMBER: 09/684,393
 ; PRIOR FILING DATE: 2000-10-10
 ; PRIOR APPLICATION NUMBER: 60/172,600
 ; PRIOR FILING DATE: 1999-12-20
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 1138
 ; TYPE: PRT
 ; ORGANISM: Rattus norvegicus
 US-10-261-482-4

Query Match 39.18; Score 1769; DB 9; Length 1138;
 Best Local Similarity 37.18; Pred. No. 4.6e-134;
 Matches 427; Conservative 103; Mismatches 254; Indels 368; Gaps 16;

QY 31 MKGCVYLGLEPL-GEAEAGLRSTRSPVCTR--FSSNGILMALAMKAVEEINN- 85
 DB 2 LRGFLAGLFLSLGDC---LQYRHRPLVYSCRPSDFNGHGHILQAMFYEELNNS 57
 QY 86 KSDLLPGLRLGYDLFTCSSEPVAMKPSLMFLARAGSHDAAYCNYQYQPRVLAVIGPH 145
 DB 58 QSDLLPGLRLGYDLFTCSSEPVAMKPSLMFLARAGSHDAAYCNYQYQPRVLAVIGPH 117
 QY 146 S----- 146
 DB 118 SSLPLGLTDCSELLAGRIVAGPSSALLPNTLLGYELDYVCSANAYATLRVLAIQPR 177
 QY 147 -----SELAMTGRKFSEFLMPQVSYGASMLLSARETFP 181
 DB 178 RHIEIQRDHRHSSKVVAFIPDNOSLAWTGRKFSEFLMPQVSYGASMLLSARETFP 237
 QY 182 SEFTVPSDRQLTAALAELOEFGW----- 206
 DB 238 SEFTVPSDRQLTAALAELOEFGWATFLMPVSYASLSAPPSFTVPSDRQLLOFGWSTD 297
 QY 207 -----N 207
 DB 298 HAVTTAALLGPFLLMPVLYSEASSVYLSAKKRPSEFLRTVPSDRHQUEVYVQLOLQSGMGN 357
 QY 208 WVAALGSDDEXGROGLSTFSALAAARGICIAHEGLVPLPRADBSRLGVQDVLHQVN--- 264
 DB 358 WVAALGSDDEXGROGLSTFSALAAARGICIAHEGLVPLPRADBSRLGVQDVLHQVNQNG 417
 QY 265 ----- 264
 DB 418 SYGLIARGICAVPRDROOSWISLIGSYGYGOLGVOALEELAVPRGICVAFKDIPEFA 477
 QY 265 -----OSSVOVLLFASVHAHALFNYSTISSRLSPRYWASEMMLTSDL 308
 DB 478 RVGDPFRMOXMOHLAOGSSVOVLLFASVHAHALFNYSTISSRLSPRYWASEMMLTSDL 537
 QY 309 VMGLPGMAOMGTVLGFQRCQAOLHEF----- 334
 DB 538 VMGLPGMAOMGTVLGF-----VFHAFIAKVMVASEMSGTYVGARTYVVVFSNRHLARVF 594
 QY 335 -----POYVKTHALATDPARCSALGEROGLLEE 363
 DB 595 FRSVYLANLTKGVVASEDNAISTYITSVIGIGITGVLAQVAAVQOQVPGLEKEEESYVR 654
 QY 364 DV-----VGORCPQCCITLQNVS--AGLNHHQFESVYTAANYVAQALAHNT 407

DB 655 AVTAAPSACPEGSWCSTNQLCRECHFTFTTRMPTLIGAFSMAAAYVEAYVAVHGHOL 714
 QY 408 LQNASGCPADDPVKPMOLLENNMTNLFVHGGCLPLRDSGNNVMEYDLKAMVQGSVR 467
 DB 715 LGCTSEIC-SRGVYPMQLAQITKYNFLHENTVARDNDGTYDITIAMDNNGEWT 773
 QY 468 LHDVGRNGS--LRTRELKIRMTSDNOKPVSRCROCGQVRYRGKGFHSCCYDCVD 524
 DB 774 FEIGSASLSPVHLDINKTKIQMGKNQVVSCTDCLAGHRRVYVGHSCFECVPC 833
 QY 525 EAGSYRONPDDIACFEGGODEWSPERSTRCRRSRLRANGEPVALLILSLALGLV 584
 DB 834 EAGFLNMSELHIQPCGTEWMAKESCTPCPRVEELAMEPISTVLIAANTLLLLLV 893
 QY 585 AALGLFVHRDSPVQASGGPLACFGLVCLVCLVSLVLFPGQSPARCLAQOPLSHLP 644
 DB 894 GTAGLFAMHFTPVYRNASGRLCFMLGSLVAGSCSYSPFGEPTVACILROPLFSLGF 953
 QY 645 TGCLSTFLQAAELFV---ESELPL---SWADRLSGCLRQPMAMVLLAMLEVALCT 697
 DB 954 AIFLSCLTIRSFOLVITFKFSTKVPTFYRTWA-----QNHGAGLFIYVSVVHLLICL 1006
 QY 698 WYLAFPPEVYTDHMLPTEALVHCRTSRWSFGLATATNATLAFLOFLGFLVRSQPC 757
 DB 1007 TWLVMMPTPTREYQRPPLVILECTEVSNGVFLAFTHTLLISTFVCSYLGEKLPEN 1066
 QY 758 YNRARGLFAMLAFTYVWSFVPLIANVQVLRPAVOMGALLICVLILAFHLPRLCYLL 817
 DB 1067 YNEAKCTYFSLILNFVSHIAFTMASTYQGSYLPAVAVYLAGLITLSGFSGLPKCTVI 1126
 QY 818 MRQGLNTPPEFF 829
 DB 1127 LCRLPRLNTEHF 1138

RESULT 15
 US-09-361-652-1
 ; Sequence 1, Application US/09361652
 ; Publication No. US20030036630A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Zuker, Charles S.
 ; APPLICANT: Lindemeyer, Juergen
 ; APPLICANT: Ryda, Nick
 ; APPLICANT: Hoon, Mark
 ; TITLE OF INVENTION: The Regents of the University of California
 ; TITLE OF INVENTION: Nucleic Acids Encoding a G-Protein Coupled Receptor
 ; FILE REFERENCE: 02307E-08861005
 ; CURRENT APPLICATION NUMBER: US/09/361,652
 ; CURRENT FILING DATE: 1999-07-27
 ; EARLIER APPLICATION NUMBER: US 60/094,465
 ; EARLIER FILING DATE: 1998-07-28
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 1
 ; LENGTH: 840
 ; TYPE: PRT
 ; ORGANISM: Rattus sp.
 ; OTHER INFORMATION: rat G-protein coupled receptor B3 (GPCR-B3)
 US-09-361-652-1

Query Match 26.7%; Score 1208.5; DB 9; Length 840;
 Best Local Similarity 33.8%; Pred. No. 5.7e-89;
 Matches 291; Conservative 151; Mismatches 339; Indels 79; Gaps 18;

QY 12 WALHPGTGAPLICSOOL-----RMKGVYLGLEPL-GEAEAGLRAS 53
 DB 4 WA-----AHLLSLQLVYCWMAFSCQRTSSPQSLPGLAGLFLSLGDC---LQY 52
 QY 54 RTRPSPVCTR--FSSNGILMALAMKAVEEINNKSDDLPLGLRLGYDLFTCSSEPVAM 110

```
Db      53  RHRPLVTSQDRPDSFNGHGYHLFOAMRFVTEELINSSALLPNTTLGLEYDYVCSSES-ANY 111
QY      111  KPSLMEFLAKAGSRDIAAYCNTOYQPRVLAVGPHSSELAAMTGKFFSFFLMPQVSYGAS 170
Db      112  YATLRVLAIOGPRHIEIQDLRNHSSKVVAFIGPDNDHAYTTALLGPFLMPLVSEAS 171
QY      171  MELLASARETFPSFFFTVPSPDRVOLTAAABLLQEFGNMVAALGSDDEYGRQCLSTIFSALA 230
Db      172  SYVLSAKRKFPSPFLRVPSPDRQVEVMQVLLQSFQWVWISLIGSYDYGQLGVALFEELA 231
QY      231  AARGICIAHEGLVPLERADDSRLG--KVQDVLHQNQSSVQVLLFASVVAHAALFNYSI 288
Db      232  VPRGICVAFKDIYPF-----SARVGDPRMQSMQHLAQARTVYVVSNNRLARVFFRSYV 287
QY      289  SSRLSPKVVAASEAMTSLDVNGLPQMAQMTYGLGLOGAQ---LHEFPQ--YKTHLAL 344
Db      288  LANLTGKVVAVASEDMATISYITSVTGIQIGYLVAVQOQVPGGLKEFEESYVRAVTA 347
QY      345  AT---DPAFCSALGEREQGLEEDVVGQRCPOCDITLQNVS--AGLNHHQTFSVYAAYS 399
Db      348  PSACPEBGSMS-----TNQLCRECHFTTRNMPTLGAFSMAATRYEAVYA 394
QY      400  VAOALHNTIOCNASGCPADPYKPMQLENNMYNLFFHVGLPLREDDSGNVDMEXDKLM 459
Db      395  VAHGLHQLLGTSEIC-SRGPIYPMOLLQIYKVNFLHENTVAFDNDGDTLGYDIIAM 453
QY      460  VMQGSVPRHLDVGRNGS---LTERLKTIRMTSDNQKPVSCSRQCEGQVRRKRGHS 516
Db      454  DNGPBMWTEIIGSASLSLVHLINKTKIQMHGKNQVPSVCTDCLAGHHRVVGSHH 513
QY      517  CCYDCVDCENAGSYRONPDIACTFCGODEMSPERSTRCFRRSRFLAAGEPAVLLLLIL 576
Db      514  CCFECVPCAGTFLNMSELHIOPCGTSEMAREKSTCTCPRTVEFLAMHPIISVLIAMN 573
QY      577  STALGLVLAALGLFVHHRDSPLVOASGGPLACGLVCLISVLLFPQSPARCIAQ 636
Db      574  TLLLLLVGTAGLFAMHRTPVVRSAGRLCFMLGLSLVAGSCSFYSFGEPTVPACLIR 633
QY      637  OPLSHLPLTGCSTLFLQAAEIFV---ESLPL---SWADRLSGCLRGPMAMIVLLAM 689
Db      634  QPLFSIGFAIFISCLTIRSFQVLIIFKSTKVPTFTRTWA-----QNHGAGLEVYBS 686
QY      690  LVEVALCTWYLVAFPPEVVTDMHMLPTALVHCRTSRWVSFGIAHATNATLAFLCLGTF 749
Db      687  FVHLICLFTLVMWTPRPTREYQRPRLVILBCTEVNSVGFLLAFTHNILLSISTFVCSY 746
QY      750  LVRSQPGCYNRARGLTFAMLAFTITWVSFVPLIANQVYLPRAVOMGALLLCVLTIAAF 809
Db      747  LKRELPEENTNEAKCVTFSLINFSWIAFTMASIYQGSYLPVAVNLAGLTLLSGGFSGY 806
QY      810  HLPKCYILMRPGPLNTPPEF 829
Db      807  FLPKCYIILCRPELNNTEHF 826
```

Search completed: May 19, 2003, 10:04:45
Job time : 43.8231 secs

THIS PAGE BLANK (USPTO)